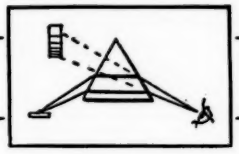
Journal of the OPTICAL SOCIETY *of* AMERICA

VOLUME 27
NO. 7



JULY
1937

Theory of Three-Color Reproduction . . . . . ARTHUR C. HARDY AND F. L. WURZBURG, JR. 227

Experimental Determination of Photographic Density . . . . . A. M. KOERNER AND CLIFTON TUTTLE 241

Refractometer for the Near Infrared . . . . . RUDOLF KINGSLAKE AND HILDA G. CONRADY 257

Light Signals on Moving Bodies as Measured by Transported Rods and Clocks . . . . . HERBERT E. IVES 263

Published monthly
at Prince and Lemon Sts., Lancaster, Pa.
by the
AMERICAN INSTITUTE OF PHYSICS
Incorporated
175 Fifth Avenue, New York, N. Y.
FOR AND UNDER THE EDITORIAL SUPERVISION OF
THE OPTICAL SOCIETY OF AMERICA

THE OPTICAL SOCIETY OF AMERICA

OFFICERS OF THE SOCIETY

A. C. HARDY, *President*
Massachusetts Institute of Technology
Cambridge, Massachusetts

R. C. GIBBS, *Vice President*
Department of Physics
Cornell University
Ithaca, New York

L. B. TUCKERMAN, *Secretary*
National Bureau of Standards
Washington, D. C.

HENRY F. KURTZ, *Treasurer*
Bausch & Lomb Optical Company
Rochester, New York

The Board of Directors consists of the above officers; the Editor of the Journal; the Past President; and five elected members.

W. B. RAYTON, *Past President*
Bausch & Lomb Optical Company
Rochester, New York

F. K. RICHTMYER, *Editor*
Cornell University
Ithaca, New York

Elected Members

ARTHUR H. COMPTON
Ryerson Physical Laboratory
University of Chicago
Chicago, Illinois

K. S. GIBSON
Bureau of Standards
Washington, D. C.

G. R. HARRISON
Department of Physics
Massachusetts Institute of Technology
Cambridge, Massachusetts

HERBERT E. IVES
463 West Street
New York, New York

CARL W. KEUFFEL
Keuffel & Esser Company
Hoboken, New Jersey

ROCHESTER, NEW YORK SECTION OF THE OPTICAL SOCIETY OF AMERICA

L. A. DUBRIDGE, *President*
F. M. E. HOLMES, *Secretary*

B. NOYES, *Vice President*
C. A. MORRISON, *Treasurer*

Councilors: F. E. ALTMAN, F. C. FAIRBANKS, L. V. FOSTER, O. SANDVIK, C. C. NITCHIE (*ex officio*)

THE OPTICAL SOCIETY OF AMERICA was organized in 1916 "*to increase and diffuse the knowledge of optics in all its branches, pure and applied, to promote the mutual interests of investigators of optical problems, of designers, manufacturers and users of optical instruments and apparatus of all kinds and to encourage cooperation among them.*" It aims to serve the interests of all who are engaged in any branch of optics from fundamental research to the manufacture of optical goods.

MEMBERSHIP

Regular Members: "Any person who has contributed materially to the advancement of optics is eligible for regular membership."

Associate Members: "Any person who is interested in optics is eligible for associate membership."

Corporate Members: "Any corporation desirous of encouraging the work of the Society is eligible for corporation membership."

The Society cordially invites to membership all who are interested in any branch of optics, either in research, in instruction, in optical or illuminating engineering, in the manufacture and distribution of optical goods of all kinds or in physiological and medical optics.

ANNUAL DUES

Regular members, $7.50; Associate members, $5.00; Corporate members, $50.00. These dues include subscription to the Society's official journal, the *Journal of the Optical Society of America.*

MEETINGS

The Society holds two meetings per year: An annual meeting in the autumn and a midwinter meeting jointly with the American Physical Society.

For further information concerning the Optical Society of America address the Secretary, Dr. L. B. Tuckerman, National Bureau of Standards, Washington, D. C.

Journal of the
OPTICAL SOCIETY *of* AMERICA

VOLUME 27 JULY, 1937 NUMBER 7

The Theory of Three-Color Reproduction

ARTHUR C. HARDY
Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts

AND

F. L. WURZBURG, JR.
Research Laboratories, Interchemical Corporation, New York, N. Y.

(Received April 21, 1937)

THE arts of color photography and color printing may be said to have resulted from an observation made by J. C. LeBlon during the early part of the 18th century. LeBlon was making prints by the mezzotint process, which consists in first tooling a copper plate until it has a uniformly rough surface. The desired design is then imposed on the plate by scraping away the burr left by the tooling operation. When the plate is inked and wiped, the original rough surface retains more ink than the parts that have been scraped. LeBlon tried making color prints, at first by printing from seven plates with inks having the hues of the seven spectral colors of Newton. Later, in 1722, he came to the conclusion that all colors can be reproduced by printing from only three plates with red, yellow, and blue inks. This recognition of the adequacy of three primary colors established the first principle underlying all methods of three color reproduction.

The second important step in the development of three-color processes was made by Clerk Maxwell in 1855. Maxwell suggested a photographic procedure in which three negatives of the subject are made, each negative being exposed through an appropriate color filter. These color-separation negatives are then used to control automatically the amounts of the three primaries at each point of the positive reproduction. All methods of three color reproduction that have since been devised, both in photography and in the graphic arts, utilize these two basic principles.

The procedure suggested by Maxwell[1] is summarized in his own words as follows: "Let it be required to ascertain the colours of a landscape by means of impressions taken on a preparation equally sensitive to rays of every colour. Let a plate of red glass be placed before the camera, and an impression taken. The positive of this will be transparent wherever the red light has been abundant in the landscape, and opaque where it has been wanting. Let it now be put in a magic lantern along with the red glass, and a red picture will be thrown on the screen. Let this operation be repeated with a green and a violet glass, and, by means of three

[1] Maxwell, Trans. Roy. Soc. Edinburgh **21**, 275 (1855).

BOSTON UNIVERSITY
COLLEGE OF LIBERAL ARTS
LIBRARY

magic lanterns, let the three images be superimposed on the screen. The colour of any point on the screen will then depend on that of the corresponding point of the landscape; and, by properly adjusting the intensities of the lights, etc., a complete copy of the landscape, as far as visible colour is concerned, will be thrown on the screen. The only apparent difference will be that the copy will be more subdued, or less pure in tint, than the original. Here, however, we have the process performed twice—first on the screen, and then on the retina."

The above quotation follows an outline of Young's theory of color vision, and the process to which Maxwell refers in the last sentence is the process that Young had assumed to take place in the human eye. Maxwell evidently sensed that the characteristics of the visual apparatus of the observer are involved in three-color reproduction. He apparently appreciated also that, even with flawless technique, the reproduction made by his method would not be perfect. For some reason he failed to recognize that the "sensation curves" which he obtained in the course of another investigation provide the fundamental basis for the theory of trichromatic reproduction. The importance of these sensation curves was not realized until attention was drawn to them by F. E. Ives. The theory proposed by Ives precipitated a controversy which raged for many years. When it was discovered that Maxwell's sensation curves are merely mixture curves for a set of three arbitrarily selected primaries, Ives' opponents erroneously concluded that his theory had been vitiated. One purpose of this paper is to show how the characteristics of the visual apparatus of the observer may be correctly embodied in a general theory of three-color reproduction and incidentally to demonstrate the inadequacy of the purely objective theory which seems to prevail at the present time.

The reasoning underlying the objective theory is apparently that, if an additive process like Maxwell's were to employ a large number of primary colors in the reproduction, the primaries might be made to consist of nearly homogeneous bands of radiant energy. In this case, each color-separation negative should be exposed through a filter that transmits the wave-length band of the corresponding primary. This means that, if filters are used in the projection of the positives, they should be the same filters that are employed in the exposure of the negatives. In the corresponding subtractive processes employing a large number of primaries, the dyes or pigments used in the preparation of the final positive image should have narrow absorption bands. Furthermore, no two absorption bands should overlap, and no gaps should exist between the bands. Each color-separation negative should in this case be responsive only to light in the spectral region absorbed by the dye whose concentration it controls. In other words, the colors of the taking filters should be complementary to the colors of the dyes employed in the positive reproduction.

Such reasoning is valid when the number of primaries is large. In fact, if it is sufficiently large, the reproduction may be made to duplicate the spectral quality of the light reflected from every area of the subject. Such a result would be considered perfect by even a color-blind observer. Practical considerations prevent the realization of such a result because of the requirement that the number of primaries be kept at a minimum. We now know that LeBlon was slightly in error when he claimed that all colors can be produced by mixtures of three primaries. The statement is rigorously correct only if negative amounts of the primaries are permitted. With this amendment, the statement is true for any set of three primaries. Again, practical considerations require that the primaries be so chosen that a satisfactory gamut of colors can be realized with mixtures containing only positive amounts of the primaries. Fortunately, the dyes and pigments that are available today enable a gamut to be selected that will include all the colors encountered in the vast majority of subjects.

When the primaries that are to be used in the reproduction have been selected, the next problem is to control the amounts of the primaries properly at every point of the reproduction. LeBlon solved this problem by making his printing plates entirely by hand, his skill and patience being the only factors limiting the perfection of the result. If it is desired to adopt Maxwell's suggestion that the process be made

automatic, the color-separation negatives have a definite function to perform. That they are not ordinarily made to perform this function properly is evident from the large amount of handwork that is still necessary in the preparation of printing plates. Where handwork is impracticable, as in motion picture photography, the desired colors are customarily obtained in the reproduction by a deliberate falsification of the colors of the subject. One reason for this situation is the adoption of techniques based on the objective theory of three-color reproduction. This theory, although valid when the number of primaries is large, is grossly inadequate in the case of a three-color process.

The Theoretical Requirements for Exact Color Reproduction

Suppose the tristimulus values of unit amounts of the primaries which are mixed to form the positive reproduction are $X_rY_rZ_r$, $X_gY_gZ_g$, and $X_bY_bZ_b$, respectively.[2] At some point of the reproduction, let the mixture consist of r units of the red primary, g of the green, and b of the blue.[3] Then, the tristimulus values of this mixture are

$$X' = rX_r + gX_g + bX_b, \tag{1a}$$

$$Y' = rY_r + gY_g + bY_b, \tag{1b}$$

$$Z' = rZ_r + gZ_g + bZ_b. \tag{1c}$$

Let the light reflected from the corresponding point of the subject have a spectral distribution of energy E. The tristimulus values at this point of the subject are

$$X = \int_0^\infty E\bar{x}d\lambda, \tag{2a}$$

$$Y = \int_0^\infty E\bar{y}d\lambda, \tag{2b}$$

$$Z = \int_0^\infty E\bar{z}d\lambda, \tag{2c}$$

where $\bar{x}$, $\bar{y}$, and $\bar{z}$ are the distribution functions of the assumed observer. Although perfect reproduction is attained theoretically only when $X' = X$, $Y' = Y$ and $Z' = Z$, the reproduction will be entirely satisfactory in practice over a wide range of brightness levels if

$$X' = kX, \tag{3a}$$

$$Y' = kY, \tag{3b}$$

$$Z' = kZ, \tag{3c}$$

where k is a proportionality constant. The problem in any automatic process is to devise a technique that will cause Eqs. (3) to be satisfied regardless of the form of the function E.

The first step in virtually all methods of three-color reproduction is the making of three color-separation negatives (or the equivalent). Let the effective spectral sensitivity of the three photographic materials be represented by S_r, S_g, and S_b, respectively, the effective sensitivity at each wave-length being the product of the sensitivity of the emulsion and the transmittance of the filter used in conjunction therewith. Let the exposures impressed on the three emulsions when acted upon by light with an energy distribution E be Σ_r, Σ_g, and Σ_b, respectively, where

$$\Sigma_r = \int_0^\infty ES_r d\lambda, \tag{4a}$$

$$\Sigma_g = \int_0^\infty ES_g d\lambda, \tag{4b}$$

$$\Sigma_b = \int_0^\infty ES_b d\lambda. \tag{4c}$$

Without making any assumptions regarding the details of the subsequent procedure, let the process be one in which r, the amount of the red primary in the reproduction, depends only on Σ_r, and is independent of Σ_g and Σ_b. If due attention is paid to the reproduction of tone values, which is a necessary precaution even when mak-

[2] The symbols and terminology here employed and the data concerning the standard observer will be found in the *Handbook of Colorimetry* (Technology Press, Massachusetts Institute of Technology, Cambridge, Massachusetts, 1936).

[3] The terms *red*, *green*, and *blue* are adopted to facilitate the presentation and do not cause any sacrifice of generality. Those who are accustomed to think in terms of subtractive processes will find it useful to keep in mind that the conventional blue-green dye or pigment controls the amount of red light in the reproduction; the magenta (red) dye or pigment, the amount of green light; and the yellow dye or pigment, the amount of blue light. A method by which the primaries in a subtractive process may be identified is discussed in a later section.

ing black and white positives, r will be directly proportional to Σ_r. By making similar assumptions with respect to the green and blue primaries, it follows that

$$r = k_r \Sigma_r = k_r \int_0^\infty E S_r d\lambda, \tag{5a}$$

$$g = k_g \Sigma_g = k_g \int_0^\infty E S_g d\lambda, \tag{5b}$$

$$b = k_b \Sigma_b = k_b \int_0^\infty E S_b d\lambda. \tag{5c}$$

These values of r, g, and b may now be substituted in Eqs. (1), and the resulting values for X', Y', and Z' may be substituted in Eqs. (3). Likewise, the values of X, Y, and Z from Eqs. (2) may be substituted in Eqs. (3). The constant k may be combined with the other constants, the resulting equations being

$$k_r X_r \int_0^\infty E S_r d\lambda + k_g X_g \int_0^\infty E S_g d\lambda + k_b X_b \int_0^\infty E S_b d\lambda = \int_0^\infty E \bar{x} d\lambda, \tag{6a}$$

$$k_r Y_r \int_0^\infty E S_r d\lambda + k_g Y_g \int_0^\infty E S_g d\lambda + k_b Y_b \int_0^\infty E S_b d\lambda = \int_0^\infty E \bar{y} d\lambda, \tag{6b}$$

$$k_r Z_r \int_0^\infty E S_r d\lambda + k_g Z_g \int_0^\infty E S_g d\lambda + k_b Z_b \int_0^\infty E S_b d\lambda = \int_0^\infty E \bar{z} d\lambda. \tag{6c}$$

Now consider the implications of these equations. If all colors are to be reproduced faithfully, Eqs. (6) must be valid regardless of the form of the function E. This will be true only if at every wave-length

$$k_r X_r S_r + k_g X_g S_g + k_b X_b S_b = \bar{x}, \tag{7a}$$

$$k_r Y_r S_r + k_g Y_g S_g + k_b Y_b S_b = \bar{y}, \tag{7b}$$

$$k_r Z_r S_r + k_g Z_g S_g + k_b Z_b S_b = \bar{z}. \tag{7c}$$

By way of review, $X_r Y_r Z_r$, $X_g Y_g Z_g$, and $X_b Y_b Z_b$ are the tristimulus values of unit amounts of the primaries; S_r, S_g, and S_b are functions representing the effective spectral sensitivities of the photographic materials used in making the color-separation negatives; and $\bar{x}$, $\bar{y}$ and $\bar{z}$ are the distribution functions of the observer in the colorimetric system employed. The constants k_r, k_g, and k_b are in practice to be assigned values that will cause one color in the subject to be reproduced correctly. When this is done, all other colors will be reproduced correctly if Eqs. (7) are satisfied.

Inspection of these equations shows that, if the tristimulus values of any primary were multiplied by a constant, it would merely alter the value of the constant associated with the primary. These equations may therefore be rewritten in terms of the trichromatic coefficients of the primaries. Since it is the form of the functions S and not the absolute values that are of interest, the constants may be discarded. The fundamental conditions for exact color reproduction then become

$$x_r S_r + x_g S_g + x_b S_b = \bar{x}, \tag{8a}$$

$$y_r S_r + y_g S_g + y_b S_b = \bar{y}, \tag{8b}$$

$$z_r S_r + z_g S_g + z_b S_b = \bar{z}. \tag{8c}$$

These equations suffice to determine the spectral sensitivity requirements that must be fulfilled by the three color-separation negatives.

Application of Theory to Additive Processes

The foregoing theory is immediately applicable to any additive process because the primaries of such a process are easily identified and their trichromatic coefficients remain constant throughout the reproduction. In the process suggested by Maxwell, for example, the trichromatic coefficients of each primary can be found by measuring the color of the projection screen with a colorimeter when the screen is illuminated by only one lantern. Such measurements depend to some extent on the observer who uses the colorimeter, and a more significant result is obtained by computations based on the standard observer recommended by the 1931 International

Commission on Illumination. In other words, the energy distribution of the light reflected by the projection screen when illuminated by each lantern can be ascertained by determining at each wave-length the product of the energy emitted by the lamp, the transmittance of the color filter, and the reflectance of the screen. With this information, the trichromatic coefficients of the primaries can be calculated by employing procedures that are well standardized in the field of colorimetry.

It has already been indicated that the primaries in a reproduction system should be chosen in such a manner that mixtures containing positive amounts of the primaries will produce a color gamut which is as wide as possible. The selection of the primaries is greatly facilitated by the use of a chromaticity diagram like that shown in Fig. 1, which is based on the I.C.I. standard observer and coordinate system. The points R, G, and B in this figure represent homogeneous radiant energy at 700, 535, and 400 millimicrons, respectively. Using such primaries in an additive process, the gamut of colors attainable[4] with positive amounts of the primaries is represented by the triangle RGB.

With the standard I.C.I. observer and coordinate system, the trichromatic coefficients of these primaries are given in Table I. These trichromatic coefficients are to be substituted in Eqs. (8) which, when solved explicitly for S_r, S_g, and S_b, become

TABLE I.

RED PRIMARY ($\lambda = 700$ mμ)	GREEN PRIMARY ($\lambda = 535$ mμ)	BLUE PRIMARY ($\lambda = 400$ mμ)
$x_r = 0.7347$	$x_g = 0.1929$	$x_b = 0.1733$
$y_r = 0.2653$	$y_g = 0.7816$	$y_b = 0.0048$
$z_r = 0.0000$	$z_g = 0.0255$	$z_b = 0.8219$

[4] Although other sets of primaries can be found that will slightly increase the area of this triangle, the set selected has the advantage of permitting the reproduction of colors of high purity with dominant wave-lengths lying between 570 and 640 millimicrons. The colors of natural objects and also the colors produced with existing dyes and pigments commonly approach the spectrum locus more closely in this spectral region than elsewhere. It should be noted that the chromaticity diagram does not permit the representation of brightness. The gamut of attainable colors is actually a solid in which brightness is represented in a direction perpendicular to the plane of the diagram. For a complete representation of the attainable gamut, this solid must be constructed.

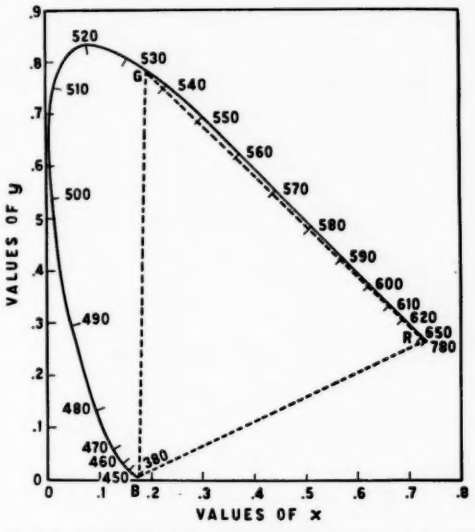


FIG. 1. Chromaticity diagram based on the I.C.I. standard observer and coordinate system. The dashed triangle indicates the color gamut obtainable with mixtures containing positive amounts of the homogeneous primaries R, G, and B, whose wave-lengths are 700, 535, and 400 millimicrons, respectively.

$$S_r = (y_g z_b - y_b z_g)\bar{x} + (x_b z_g - x_g z_b)\bar{y} + (x_g y_b - x_b y_g)\bar{z}, \quad (9a)$$

$$S_g = (y_b z_r - y_r z_b)\bar{x} + (x_r z_b - x_b z_r)\bar{y} + (x_b y_r - x_r y_b)\bar{z}, \quad (9b)$$

$$S_b = (y_r z_g - y_g z_r)\bar{x} + (x_g z_r - x_r z_g)\bar{y} + (x_r y_g - x_g y_r)\bar{z}, \quad (9c)$$

after eliminating a proportionality constant which is unnecessary because only relative values of the functions S_r, S_g, and S_b are required. With the aid of a table giving the values of $\bar{x}$, $\bar{y}$, and $\bar{z}$ at convenient intervals of wave-length, values for S_r, S_g, and S_b can be found at the corresponding wave-lengths. The results of these operations in the case of the assumed homogeneous primaries are shown graphically in Fig. 2, the ordinates representing on an arbitrary scale the effective spectral sensitivities of the three negative emulsions required for the proper control of these primaries, the abscissae representing wave-length. In the event that the photographic emulsions employed are uniformly sensitive to all wave-lengths, these curves represent the transmission characteristics of the required taking filters.

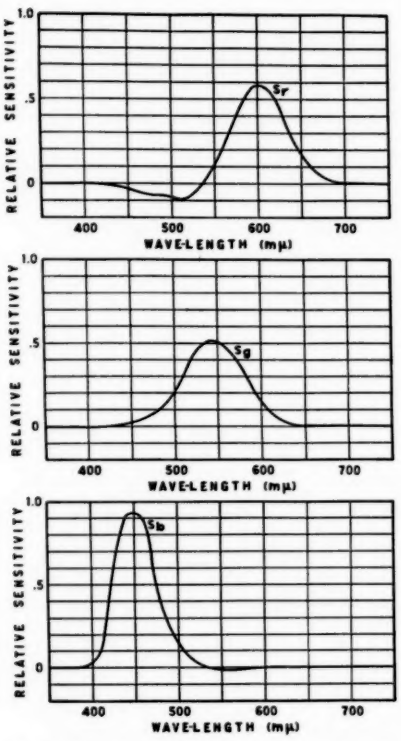


FIG. 2. The functions S_r, S_g, and S_b represent the effective spectral sensitivities of the three emulsions to be used in making three color-separation negatives that will control properly the homogeneous primaries represented in Fig. 1.

It will be instantly recognized that these curves bear no more than a slight resemblance to the effective spectral sensitivity characteristics of the photographic materials that are commonly used or recommended for the production of color-separation negatives. In particular, these requirements demand that the sensitivity be less than zero in certain spectral regions. Such negative values of the sensitivity can evidently not be achieved with color filters which, in the limit, can merely reduce the effective sensitivity to zero.

It is evident from the manner of deriving the sensitivity curves of Fig. 2 that their form is identical with the color mixture curves for this set of primaries. These curves differ from Maxwell's "sensation curves" chiefly because the latter represent color mixture data for a different set of primaries. Maxwell's curves possess negative values in certain spectral regions, and it has long been recognized that mixture curves containing no negative values are possible only when the primaries are imaginary. Most authors who have developed the trichromatic theory of color reproduction have either adopted color mixture curves based on imaginary primaries, have ignored the negative portions of curves based on real primaries, or have taken pains to point out that, because of the negative values, such curves are merely of theoretical interest.[5] In later sections of this paper an attempt will be made to indicate how these color mixture curves may be given a real significance, and also how this theory is applicable to subtractive processes which, although of more practical importance, have not been subjected to a trichromatic analysis.

It will be useful later to have available the spectral sensitivity curves required in the case of a set of nonhomogeneous primaries. Suppose these nonhomogeneous primaries to consist of three broad contiguous spectral bands instead of the three spectral lines previously assumed. More specifically, suppose that these bands are obtained by isolating three broad spectral regions from the energy emitted by the I. C. I. illuminant *C*, a standard source representative of daylight. If the boundaries between the contiguous bands are arbitrarily placed at 490 millimicrons and 580 millimicrons, the relative spectral distribution of energy in the three primaries will be as illustrated in Fig. 3. The trichromatic coefficients of these primaries are found by the application of standardized procedures to be as shown in Table II.

The color gamut obtainable with mixtures containing positive amounts of these primaries is represented by the triangle *RGB* in Fig. 4. This

TABLE II.

RED PRIMARY (580 mμ–780 mμ)	GREEN PRIMARY (490 mμ–580 mμ)	BLUE PRIMARY (380 mμ–490 mμ)
$x_r=0.6429$	$x_g=0.2653$	$x_b=0.1449$
$y_r=0.3567$	$y_g=0.6533$	$y_b=0.0397$
$z_r=0.0004$	$z_g=0.0813$	$z_b=0.8154$

[5] Reference should be made in this connection to the early papers of F. E. Ives and the more recent work of Schaefer and Ackermann, Zeits. f. tech. Physik **8**, 55–62 (1927).

gamut is considerably less than that shown in Fig. 1, but it is still sufficient for all ordinary purposes. For example, there are no colors in the *Munsell Book of Color* that are not included within the gamut of Fig. 4. If the trichromatic coefficients of these nonhomogeneous primaries are substituted in Eqs. (9), the functions S_r, S_g, and S_b are found to take the forms shown in Fig. 5. It will be noticed that the negative regions are of even greater importance than in the case of the functions shown in Fig. 2. This is a direct consequence of the fact that the spectrum locus lies so far outside the triangle *RGB* in Fig. 4.

The Realization of Receptors Having the Required Spectral Sensitivity

It has already been mentioned that filters are incapable of imparting to a photographic emulsion a spectral sensitivity less than zero. A negative sensitivity in certain spectral regions means that the effect of light in these regions must be to destroy rather than to create a latent image. Herschel, as long ago as 1840, discovered that the exposure of photographic emulsions to infrared light destroyed the latent image that had resulted from exposure to visible light. Some work has since been done on the Herschel effect exhibited by densensitizing dyes which cause a destruction of the latent image when the emulsion is exposed to radiant energy that is absorbed by the dye. By using desensitizing dyes that absorb in the proper spectral region, it seems probable that a very elegant solution to the problem of obtaining the desired negative sensitivity could be effected.

There are many procedures which achieve the desired result without the use of receptors having spectral sensitivities less than zero. One procedure that will be mentioned first because of its relative simplicity is useful when it is desired to transmit colored images by wire or by radio. Let the subject be "scanned" by photoelectric cells whose responses are used to control the mixture of the three primaries at each point of the image reproduced at the receiving station. Assume that the choice of primaries requires response curves like those in Fig. 5; and, for the moment, consider only the curve for the receptor

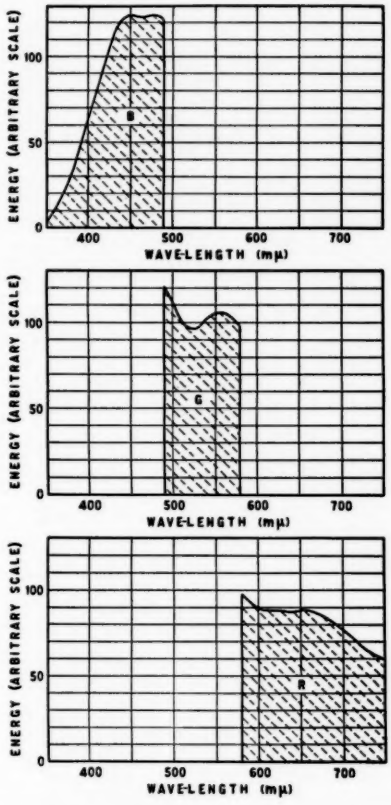


FIG. 3. Relative spectral energy distributions of three arbitrarily selected nonhomogeneous primaries.

controlling the red primary. If the negative portions of this curve are ignored, a filter or combination of filters can readily be prepared that will impart the desired positive sensitivity to a photoelectric cell. Similarly, another filter or combination of filters can be prepared that will impart to another photoelectric cell a sensitivity which is the mirror image of the portions of the curve which were previously ignored. The effective spectral sensitivities of the two photoelectric cells are shown by curves *A* and *B* in Fig. 6. If both cells are allowed to scan the same area of the subject simultaneously, a signal which will control the red primary properly at the receiving station is obtained by subtracting the response of cell *B* from that of cell *A*. This difference between the two photoelectric currents can be

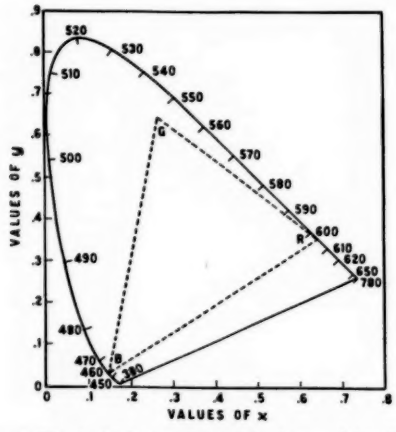


FIG. 4. The dashed triangle on this chromaticity diagram indicates the color gamut obtainable with mixtures containing positive amounts of the nonhomogeneous primaries R, G, and B, whose energy distributions are represented in Fig. 3.

obtained electrically in a variety of ways. By following a similar procedure in the case of the other receptors, it is clear that the three primaries at the receiving station can be controlled properly by using six photoelectric cells.

A similar artifice can be employed in a photographic process. In conformity with the example just cited, suppose that two negatives of the subject are made, the exposures being made through filters that result in effective spectral sensitivities like those represented by curves A and B of Fig. 6. These negatives can be used to control the red primary in the reproduction if a technique is employed which causes the amount of the primary to depend on the difference between the exposures impressed on negatives A and B. To secure such a result is more difficult than with photoelectric receptors because photographic emulsions do not ordinarily exhibit a linear response. A conventional characteristic curve for a typical photographic emulsion is shown in Fig. 7. This curve is obtained by plotting the optical density of the developed image as ordinates, and the logarithm of the exposure as abscissae. As photography is ordinarily practiced, an effort is made to adjust the time of exposure in such a manner that all the tones of the subject will be recorded on the straight portion of the characteristic curve. The equation for the straight portion of the curve is

$$D=\gamma(\log \Sigma-\log i), \tag{10}$$

where γ is a constant that depends on the extent of development, and i is the exposure at the point where the extrapolated straight line intersects the axis of abscissae.

By definition, the transmittance of a developed photographic deposit is the ratio of the light that it transmits to the light that is incident upon it. The transmittance is related to the optical density by the equation

$$D=\log (1/T). \tag{11}$$

By eliminating D from Eqs. (10) and (11), it is found that

$$T=(i/\Sigma)^{\gamma}. \tag{12}$$

When a black and white positive is made by exposure to light transmitted by the negative,

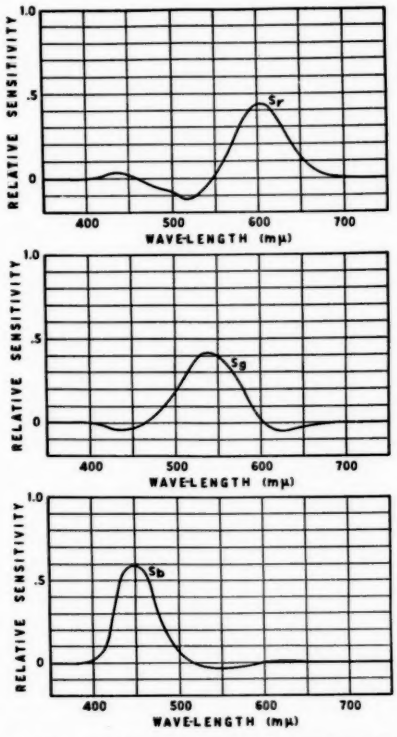


FIG. 5. The functions S_r, S_g, and S_b represent the effective spectral sensitivities of the three emulsions to be used in making three color-separation negatives that will control properly the nonhomogeneous primaries whose energy distributions are represented in Fig. 3.

the exposure at each point of the positive is proportional to the transmittance at the corresponding point of the negative. In other words,

$$\Sigma' = cT. \tag{13}$$

Since the characteristics of the positive material are in general somewhat similar to those of the negative material,

$$T' = (i'/\Sigma')^{\gamma'}. \tag{14}$$

If an image of the positive is projected on a screen, the brightness at each point of the screen is proportional to the transmittance T' at the corresponding point of the positive. The conditions for correct tone rendering are satisfied if T' is proportional to Σ. Eliminating T and Σ' from Eqs. (12), (13), and (14), it is seen that

$$T' = (i'/c(i/\Sigma)^{\gamma})^{\gamma'} = k\Sigma^{\gamma\cdot\gamma'}. \tag{15}$$

The tone rendering is evidently correct if $\gamma\cdot\gamma' = 1$. It should be noted that the relationship

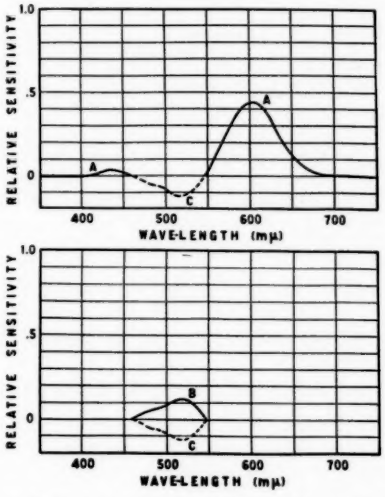


FIG. 6. Curves A and C together are identical with the S_r function shown in Fig. 5. The ordinates of curve B are the same as the corresponding ordinates of curve C with the signs changed.

between the transmittance and the exposure of the negative is hyperbolic. In the case of a gamma-product of unity, an over-all compensation is effected by the hyperbolic response of the positive material.

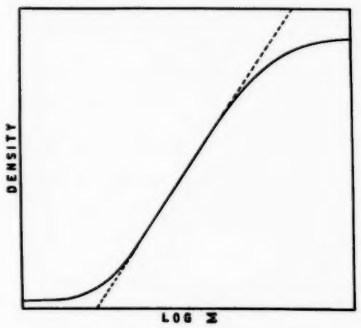


FIG. 7. Conventional characteristic curve for a typical photographic emulsion showing the density of the developed deposit as a function of the logarithm of the exposure.

Consider now a technique for obtaining correct tone rendering which is unorthodox[6] but which is susceptible of being controlled in such a manner that the desired effect of subtraction can be obtained. Referring again to Fig. 7, the region to the left of the straight portion of the characteristic curve is commonly called the region of underexposure or "toe" of the curve. If the transmittance of the developed deposit is plotted as a function of exposure in this region, a nearly linear relationship is observed. As shown in Fig. 8, some special types of photographic emulsions may be so manipulated in processing as to extend the linear portion of the exposure scale considerably. The analytical expression for the straight line which the curve of Fig. 8 approximates is

$$T = T_0 - k\Sigma, \tag{16}$$

where T_0 is the transmittance corresponding to zero exposure. If a positive is made using a material with similar characteristics,

$$T' = T_0 - k\Sigma'. \tag{17}$$

Now the exposure $\Sigma' = cT$ as before. Hence, by eliminating T and Σ' from the above equations,

$$T' = T_0(1 - ck) + ck^2\Sigma. \tag{18}$$

It appears that the transmittance T' of the positive will be proportional to the exposure Σ of the negative if the procedure is conducted in

[6] L. A. Jones, "Photographic Reproduction of Tone," J. Opt. Soc. Am. **5**, 232 (1921).

such a manner that $ck=1$. In this event,

$$T'=k\Sigma. \tag{19}$$

Now let this method of securing tone reproduction be employed to achieve the proper

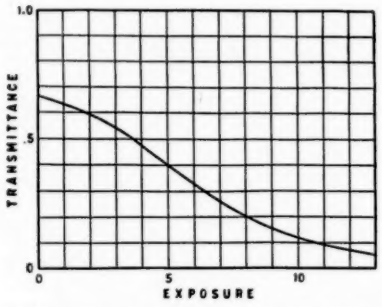


FIG. 8. Curve showing the transmittance of the developed deposit as a function of the exposure in the "underexposure" region.

control of the red primary in a three color process. Suppose two negatives of the subject are made using materials having effective spectral sensitivities corresponding respectively to curves A and B of Fig. 6. Assuming still the process of Maxwell for the purpose of illustration, the transmittance of the positive used in the projection of the red primary is related to the exposure that it receives by the expression

$$T'=T_0-k\Sigma'. \tag{20}$$

Suppose that the exposure Σ' is the sum of two exposures, one being made to light transmitted by negative A and the other to light transmitted by a positive made from negative B. In other words, let

$$\Sigma'=c(T_A+T_B'). \tag{21}$$

If the characteristics of negative A are represented by

$$T_A=T_0-k\Sigma_A, \tag{22}$$

and if the characteristics of positive B are represented by

$$T_B'=k\Sigma_B, \tag{23}$$

it follows that

$$T'=(1-ck)T_0+ck^2(\Sigma_A-\Sigma_B). \tag{24}$$

If the previous assumption is made that $ck=1$, then

$$T'=k(\Sigma_A-\Sigma_B). \tag{25}$$

In other words, the transmittance of the final positive is proportional to the difference in the exposures of negatives A and B. By following a similar procedure, the green and blue primaries may also be controlled properly. In the general case, this will require the making of six color-separation negatives, two for each primary.

The various techniques that are employed in color reproduction have been devised largely on the basis of empirical observations. The usual procedure is to reproduce a test subject containing an assortment of colored areas. An attempt is then made to deduce the proper alterations in the technique from visual estimates of the errors in the reproduction. It is not at all surprising that many methods of color reproduction have been proposed in which one negative is "corrected" by a positive made from another negative. In the so-called "masking" methods, a negative is combined in register with a thin positive made from another negative, and the light transmitted by the combination is used to control the exposure of a composite image. It is easily seen that this procedure does not produce the desired effect of subtraction. It is likewise clear that, if a composite image is made by the addition of two exposures as suggested above, the effect of subtraction is secured only if materials having an approximately linear response are employed. Many other methods of securing the effect of subtraction are possible, and the above method is cited merely as an example. It should be mentioned in passing that, in any procedure resulting in the effective subtraction of two exposures, the minuend will exceed the subtrahend for all colors of the subject that can be reproduced with the given set of primaries.

It has been shown that the effect of subtraction can be obtained either by photoelectric or by photographic means. A combination of both procedures is also possible and should be especially useful in the graphic arts. Suppose that two negatives of a subject are made using photographic materials with spectral sensitivities corresponding to curves A and B of Fig. 6. Let positives be made from these which conform to either method of satisfying the condi-

tions for correct tone rendering. These positives may then be simultaneously scanned by two photoelectric cells using methods which have been developed for the transmission of pictures by wire. The current through the cell scanning positive B may be subtracted electrically from the current through the cell scanning positive A. The net current may then be used to control a light valve (or similar mechanism) to prepare either a negative or positive image as the process being employed demands. At this stage, corrections may be introduced into the electrical circuit or light path that will compensate for nonlinearity in the subsequent steps of the process. In this way a set of six color-separation negatives can be used to control the three primaries of the reproduction with theoretical exactitude.

The necessity for six color-separation negatives constitutes a serious obstacle except when the subject is stationary. Let us therefore consider a technique that requires the making of only three color-separation negatives. This technique, although of more general applicability, will be illustrated in terms of the procedure just outlined. Let three color-separation negatives be made, and let a positive be made from each in such a manner that the transmittance of each positive is proportional to the exposure received by the corresponding negative. Then let a scanning apparatus containing three photoelectric cells be so arranged that each cell scans one of these positives. If the three positives are distinguished by subscripts 1, 2, and 3, the photoelectric current resulting from the excitation of the three photoelectric cells is given by the expression

$$I=k_1\Sigma_1+k_2\Sigma_2+k_3\Sigma_3.$$

Each constant can be made positive or negative depending upon the polarity of the photoelectric cell, and the value of each constant can be altered either optically or electrically. Let us assume that this current I is used to control the amount of the red primary at each point of the reproduction. The three positives are then scanned again using different values of the constants, and the net photoelectric current is used to control the green primary. In a similar manner, the net photoelectric current obtained with a third set of values for the constants is used to control the blue primary. We may therefore write that

$$r=k_1\Sigma_1+k_2\Sigma_2+k_3\Sigma_3, \tag{26a}$$

$$g=k_4\Sigma_1+k_5\Sigma_2+k_6\Sigma_3, \tag{26b}$$

$$b=k_7\Sigma_1+k_8\Sigma_2+k_9\Sigma_3, \tag{26c}$$

where r, g, and b are the amounts of the three primaries in the reproduction. These equations are analogous to Eqs. (5), which were derived on the assumption that each primary is controlled by a single color-separation negative. By algebraic operations similar to those employed previously, it can be shown that

$$K_1S_1+K_2S_2+K_3S_3=\bar{x}, \tag{27a}$$

$$K_4S_1+K_5S_2+K_6S_3=\bar{y}, \tag{27b}$$

$$K_7S_1+K_8S_2+K_9S_3=\bar{z}, \tag{27c}$$

where

$$K_1=k_1X_r+k_4X_g+k_7X_b,$$

$$K_2=k_2X_r+k_5X_g+k_8X_b,$$

$$K_3=k_3X_r+k_6X_g+k_9X_b,$$

$$K_4=k_1Y_r+k_4Y_g+k_7Y_b,$$

$$K_5=k_2Y_r+k_5Y_g+k_8Y_b,$$

$$K_6=k_3Y_r+k_6Y_g+k_9Y_b,$$

$$K_7=k_1Z_r+k_4Z_g+k_7Z_b,$$

$$K_8=k_2Z_r+k_5Z_g+k_8Z_b,$$

$$K_9=k_3Z_r+k_6Z_g+k_9Z_b.$$

Eqs. (27) correspond to Eqs. (7) obtained under the previous assumption.

It is evident from Eqs. (27) that the functions S_1, S_2, and S_3 can be any any linear transformations of the observer's distribution functions $\bar{x}$, $\bar{y}$, and $\bar{z}$. It is therefore possible to assign values to the constants K_1, $K_2 \cdots K_9$ that will result in values of S_1, S_2, and S_3 that are never less than zero. This means that the desired spectral sensitivities can be obtained with color filters. Since there is considerable latitude in the choice of values for the constants, the set of sensitivity functions can be selected on the basis of convenience. For example, functions can be selected which are easy to realize with conventional emulsions and filter dyes.

It should be noted that the constants K_1, $K_2 \cdots K_9$ can be selected without knowing what primaries are to be used in the reproduction.

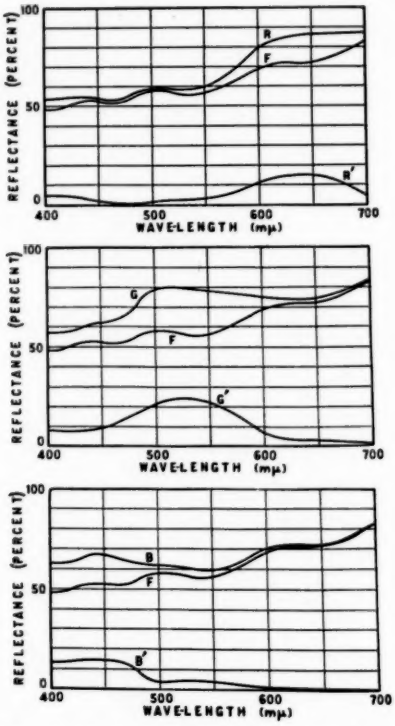


FIG. 9. Spectrophotometric curves illustrating a method for identifying the primaries in subtractive processes.

The choice of primaries then influences only the values of the constants k_1, $k_2 \cdots k_9$. Hence a single set of color-separation negatives can be used to control properly any set of primaries by merely changing the values of the constants.[7] The values of the constants may be computed, or they may be determined experimentally by the use of test charts.

APPLICATION OF THEORY TO SUBTRACTIVE PROCESSES

The traditional division of this subject into additive processes and subtractive processes is unfortunate because it exaggerates the importance of superficial details and obscures the fact that the same requirements must be satisfied by all three-color processes if faithful reproduction is to be achieved. Because the literature of this subject contains so many conflicting statements, the discussion up to this point has been restricted to a consideration of additive processes. To utilize the results of the preceding theory in the case of subtractive processes, it is now merely necessary to find a method by which the primaries in such processes can be identified.

The principle underlying the method of identifying primaries in a subtractive process will be clear from one example selected from the field of the graphic arts. Curve F in Fig. 9 is the spectrophotometric curve of a light flesh tint obtained by printing from three halftone plates with the conventional yellow, magenta, and blue-green printing inks. Curve R is a similar curve for a color produced under identical conditions except that in this case the dot size of the blue-green plate was slightly reduced. This means that somewhat less blue-green ink was transferred to the paper. It is evident that decreasing the amount of blue-green ink has increased the reflectance of the printed surface principally in the red region of the spectrum. If the surface is illuminated by an equal energy source, the curve R', whose ordinates are the differences between the ordinates of curves R and F, represents the spectral quality of the light that is controlled by the blue-green ink. In other words, it represents the spectral quality of the red primary. The trichromatic coefficients of this primary can be computed by employing any of the standardized procedures.

In a similar manner, curve G is the spectrophotometric curve of the printed surface obtained under conditions corresponding to curve F, except that in this case less magenta ink has been transferred to the paper. As before the curve G', whose ordinates are the differences between curves G and F, represents the spectral quality of the light that is controlled by the magenta ink. The trichromatic coefficients computed on the basis of curve G' are the trichromatic coefficients of the green primary under equal energy illumination. From curve B', which is obtained by diminishing the amount of yellow ink, the trichromatic coefficients of the blue primary may be determined. In the case of an illuminant other than an equal energy source, it is necessary merely to multiply the ordinates

[7] The tristimulus values of the primaries and the values of K_1, $K_2 \cdots K_9$ uniquely determine k_1, $k_2 \cdots k_9$. If desired, the functions S_1, S_2, and S_3 might be chosen in such a manner that some of the constants k_1, $k_2 \cdots k_9$ are substantially zero.

of curves R', G', and B' by the energy distribution curve of the source before performing the three integrations.[8]

If curves R', G', and B' in Fig. 9 are integrated after multiplication by the energy distribution of illuminant C, the trichromatic coefficients of the primaries are found to be those shown in Table III.

By substituting these values in Eqs. (9), the relative sensitivities of the three receptors that are required to control these primaries are found to have the forms shown in Fig. 10. It will be noted that these curves differ considerably from those obtained previously (Fig. 2 and Fig. 5), especially in respect to the relative magnitudes of the negative portions of the curves.

In the above example, the primaries were determined by making slight variations in the quantities of the inks required to produce a light flesh tint. If some other color had been chosen as a point of reference, a somewhat different set of primaries would in general have been found. It will be recalled that, in all strictly additive processess, the trichromatic coefficients of the primaries remain constant throughout the reproduction. The primaries in an additive process will therefore be called *stable*. Correspondingly, the primaries in a subtractive process will be called *unstable* when their trichromatic coefficients do not remain constant. Since the primaries in any process must be stable before automatic control of the primaries throughout the reproduction becomes feasible, it follows that, if the primaries are unstable, the reproduction can be exact only in the immediate neighborhood of the color that is used as a point of reference for the determination of the primaries. This means that stability of the primaries is an important characteristic of the dyes or pigments used in subtractive processes. It will be seen that there are, therefore, two major requirements that these dyes or pigments should fulfill if automatic color reproduction is to be achieved for a wide range of colors. First, the dyes or pigments must be capable of producing the desired color gamut, and second, they should result in primaries which are stable over this gamut.

As a guide to the selection of dyes or pigments for subtractive processes, it is useful to consider a set that would result in three stable primaries. If each dye is considered individually, its

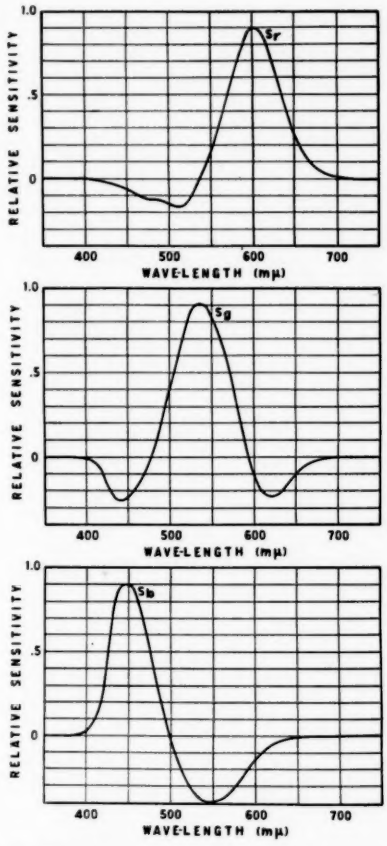


FIG. 10. The functions S_r, S_g, and S_b represent the effective spectral sensitivities of the three emulsions to be used in making three color-separation negatives that will control properly the primaries identified in Fig. 9.

TABLE III.

RED PRIMARY	GREEN PRIMARY	BLUE PRIMARY
$x_r=0.4969$	$x_g=0.1985$	$x_b=0.1847$
$y_r=0.3346$	$y_g=0.4423$	$y_b=0.1265$
$z_r=0.1685$	$z_g=0.3592$	$z_b=0.6889$

[8] An alternative procedure is to determine the trichromatic coefficients by means of a colorimeter. In any event, the preparation of these test areas should be carried out under precisely the same conditions as those obtaining in the actual process of reproduction. This means that the test areas should be printed on the same paper stock, using the inks in the same order of progression. If the process is one in which a black plate is employed, the black impression should be made in preparation of each test area.

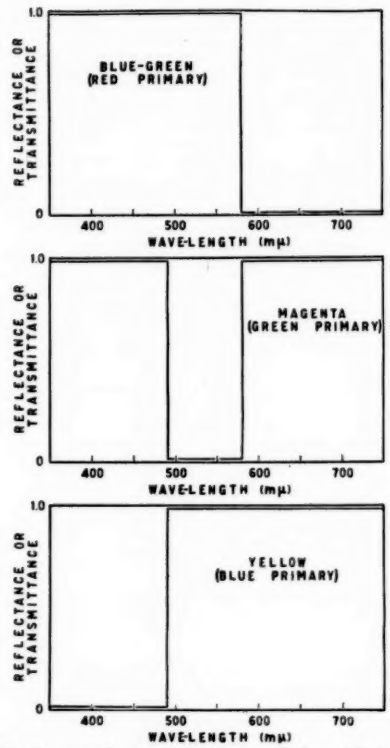


FIG. 11. Spectrophotometric curves of three ideal dyes or pigments at full concentration.

spectrophotometric characteristics must evidently be such that changes in the dye concentration (or the equivalent) will not change the trichromatic coefficients of the corresponding primary. This means that, at every wave-length, the absorption coefficient of the dye must be either zero or a constant. Stability of a set of three primaries requires in addition that, at every wave-length, the absorption coefficients of at least two of the dyes be zero. Furthermore, if the three dyes in full concentration are to produce black, the absorption coefficient of at least one of the dyes must be finite at every wave-length. Such reasoning leads to spectrophotometric characteristics of the type represented in Fig. 11. Here the blue-green dye or pigment absorbs uniformly the radiant energy in the visible region whose wave-length is longer than 580 millimicrons, the magenta dye or pigment has an absorption band between 490 and 580 millimicrons, and the yellow dye or pigment absorbs all the radiant energy in the visible region whose wave-length is less than 490 millimicrons. The assignment of 490 millimicrons and 580 millimicrons as the boundaries between the contiguous absorption bands is entirely arbitrary and is made to conform with the example cited under the discussion of additive processes.

With such an ideal set of dyes or pigments, the primaries are identical with those of Fig. 3. In this case the subtractive process becomes virtually identical with an additive process. The primaries are stable, and the spectral sensitivities which the color-separation negatives must possess are those given in Fig. 5 for the corresponding additive process.

Experimental Determination of Photographic Density*

A. M. KOERNER AND CLIFTON TUTTLE
Eastman Kodak Company, Rochester, New York

(Received February 24, 1937)

The fact that various instruments used to determine photographic density yield different results makes it desirable that some one method should be chosen as a standard. This method should produce results which are in keeping with the implications of the Hurter and Driffield definition of photographic density. A study has been made, extending some of the previous work of Tuttle, of the factors responsible for discrepancies between the readings of existing densitometers, and it is concluded that the integrating sphere densitometer is best suited to serve as a standard instrument. The relations existing between its readings and those of other densitometers are given and it it is shown that these relations are valid for a wide range of photographic materials of varying grain size. Toy's experiments to determine contact printing densities have been repeated with somewhat different results. Certain effects connected with opal densitometers and reported by Bull and Cartwright have been explained. The dependence of photographic density upon orientation of the photographic sample, that is, whether the emulsion side or base side acts as first surface, is shown. Thin metallic films on glass have been suggested to serve as standard densities. It is shown that such films are not suitable for this purpose since their optical properties are quite different from those of the photographic density.

1. THE HURTER AND DRIFFIELD DEFINITION

PHOTOGRAPHIC density was originally defined by Hurter and Driffield[1] in the following manner:

$$D = \log_{10} O = \log_{10} F_0/F_1,$$

where O, which is the ratio of F_0, the luminous flux incident on the silver deposit, to F_1, the luminous flux transmitted by the silver deposit, is termed the opacity. From the definition it is apparent that, in order to measure density, it is necessary only to determine the flux incident on, and the flux transmitted by, the silver deposit. Since there are no qualifying statements attached to the above definition, it would seem quite beside the point to inquire what portion of the fluxes should be measured; one naturally supposes that *all* the incident and *all* the transmitted flux is meant. Any other interpretation would add an equivocal quality to the term "photographic density" as defining a basic physical quantity, and although it is true that Hurter and Driffield were not aware of certain properties of the photographic image, particularly the light-scattering properties, at the time the definition was formulated, it is also probably true that, had they appreciated these factors, their definition would still have remained as it now is.

For the one who wishes to determine the density of a photographic image, as implied in the above equation, the necessary means for measuring the whole of the fluxes is available in the form of the integrating sphere. This device comprises simply a hollow, thin-walled sphere into whose periphery have been cut a pair of windows, located at approximately 90° from each other, and whose interior has been surfaced with some highly reflective and highly diffusive material such as magnesium oxide. Dimensions are immaterial, so long as the area of the apertures remains a negligible fraction of the whole surface of the sphere. It has been shown in the literature[2] that the brightness of the interior of the walls of such a sphere is directly proportional to the amount of flux entering through the aperture, regardless of direction. Hence, if light from an external source is allowed to enter through one of the openings, the flux incident at the opening can be determined by observing, through the second aperture, the brightness of the wall. If now a photographic image is laid against the first aperture and a second observation is made, all of the flux trans-

* Communication No. 612 from the Kodak Research Laboratories.

[1] F. Hurter and V. C. Driffield, "Photochemical Investigations and a New Method of Determination of the Sensitiveness of Photographic Plates," J. Soc. Chem. Ind. 9, 455 (1890).

[2] Enoch Karrer, "Use of the Ulbricht Sphere in Measuring Reflection and Transmission Factors," J. Opt. Soc. Am. 5, 96 (1921).

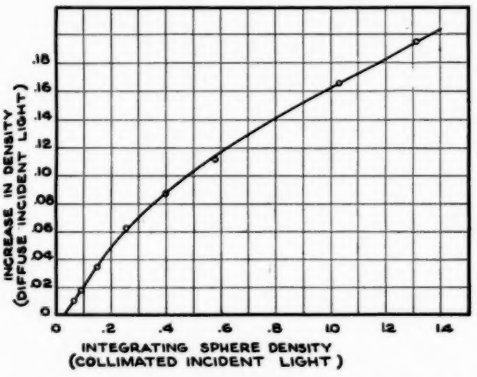


FIG. 1. Dependence of readings of integrating sphere densitometer upon diffuseness of light source.

mitted, either specularly or in a scattered manner, by the image will have been determined, and by applying the Hurter and Driffield equation, a value for the density of the image can be arrived at.

However, before density as determined in the above manner becomes a perfectly definite quantity, a certain restriction must be placed on the use of the apparatus, for, although the integrating sphere measures all the incident light and all the transmitted light, it is still possible to obtain therewith two widely different values for density, depending upon whether the incident flux, F_0, is collimated perpendicularly to the photographic image or thoroughly diffused. That this is true is shown by the curve of Fig. 1, which represents the results of measurements made exactly as described above, with first a collimated beam of light and then diffused light. Since, for the time being, the discussion is to remain general, no details of the particular instrument with which these data were determined will be presented here. Further on it will be specified completely.

It will be noted that the reading obtained with diffuse incident light is always higher than the one obtained with the collimated beam. The reason for this becomes evident when it is realized that the Hurter and Driffield equation is in reality a form of statement of Beer's law for a material of constant thickness. When diffused incident light is used, some of the rays travel through the image obliquely and are much more attenuated than those passing through normally, so that the result for these rays is the same as if the collimated beam had been used with a somewhat thicker photographic image.

The length of path traversed by a ray passing through a material bounded by plane parallel sides is given by:

$$l = nt/(n^2 - \sin^2\theta)^{\frac{1}{2}},$$

where t = thickness of the material,
n = refractive index of the material,
and θ = angle of incidence of the ray.

Numerical solution of this function, assuming an index of refraction of 1.55 for the gelatin matrix of the photographic image shows that, up to 10°, the value of l remains within 0.7 percent of the value of t. Now, if it is assumed that illumination of the film is accomplished by means of an approximate point source of light, and if the incident rays are all contained within a solid cone angle of 20°, illumination within the cone being uniformly distributed, as is approximately true in most practical optical arrangements, then the *average* value of l will deviate from t by not more than 0.25 percent and the thickness of material may be considered as being constant at the value t for all the incident rays. For greater cone angles, a thickness correction should perhaps be made. Such a correction becomes experimentally difficult with practically achievable diffuse sources, or with sources which depart from perfect collimation by varying amounts.

It is proposed by the authors that the integrating sphere used with an incident light, no ray of which departs from normal incidence by more than 10°, shall be considered as the standard optical set-up for determination of photographic density according to both letter and spirit of the Hurter and Driffield definition.[3]

2. The Integrating Sphere and Density Effective in Contact Printing

Having arrived, in a somewhat arbitrary manner, at what appears to be a reasonable basis for a system of densitometry, certain questions

[3] Experimental realization of this condition is not difficult. Even if a very efficient condenser system is used to obtain high flux density at the sphere window, the extreme rays within the entering cone will seldom exceed this value. Assume, for example, an $f:2$ condenser lens working at 1 : 1 magnification of the source. The effective aperture is $f:4$ and the cone half angle is about 9°20′.

of practicability must be considered. In the first place, since the contact print is the ultimate goal of most of the photographic processes, a definite and specifiable relationship should exist between the density effective in contact printing and density as measured by the practical system of densitometry. What we wish to determine, then, in this section, is the relation between density, as implied by the Hurter and Driffield equation and as measured by the integrating sphere, and density effective in producing a contact print.

The method for evaluating density by means of the sphere has been outlined. It remains to devise a method for determining contact printing density. This can be done simply by the application of the methods of photographic photometry.

It is quite obvious that the shades of light and dark appearing in a contact print are, in themselves, unsatisfactory as a means of determining the corresponding densities in the negative. This is true simply because they are largely dependent upon such additional factors as time and temperature of development and the characteristics of the printing medium. If, however, a sample of a printing medium is exposed through a negative to a light source, and if then a series of calibrating exposures, each of a different, but known, value, are made on the same sample, then the amount of light transmitted by the various parts of the negative can be determined by comparing the image density of those portions of the contact print with that of the portions which were exposed to the light sources of known intensity. Not only can the amount of light *transmitted by* the negative be determined in this way, but if a portion of the printing material is left uncovered during exposure of the negative, a value for the amount of light *incident upon* the negative can also be obtained. Knowing now, the amount of light incident upon, and the amount of light transmitted by, the negative, as recorded by the print, the printing density is obtained simply by applying the Hurter and Driffield equation.

Certain precautions must be observed in making the calibrating exposures. Since the photographic emulsion is not equally sensitive to all colors, the spectral quality of the light used must be identical with that of the unknown source. Also, the time of exposure for each of the calibrating exposures must be identical with the time of exposure for the unknown source. Only the intensity must be varied. This precaution eliminates errors which might enter through the failure of the reciprocity law.

In the entire literature of densitometry there is, to our knowledge, only one instance where so-called printing density, as evaluated by a contact print, has been determined without violating some of the rules of photographic photometry. This was done by Toy for several photographic materials.[4] The results thus obtained were plotted against readings obtained on a pot opal densitometer. In this manner Toy arrived at the linear relation:

$$D_{pr}=0.92D_{po},$$

where D_{pr} represents contact printing density and D_{po} represents the corresponding pot opal readings. The same relation was found to hold with good accuracy for images in several different photographic materials and on several printing media. Ferguson, using the identical photographic images employed by Toy, but without redetermining their printing densities, arrived at the relation:

$$D_{pr}=0.95D_{po}$$

simply by using a different opal densitometer.[5] Since the literature contains the results of only a single determination of contact printing densities, the experiment was repeated by us, using a slightly different apparatus as well as photographic images in different emulsions and on other bases.

The apparatus used in the experiment to be reported here is shown diagrammatically in Fig. 2. Light from the lamp filament (A) is gathered by the glass block (B), passes through the opal diffuser (K), and is then divided into two beams, one of which passes through the logarithmic aperture (C) and is then projected upon a rectangular aperture (G) in the front of a printing frame (F) by means of a cylindrical lens (E). The second beam passes through the aperture (D) which is shaped so that the illumination upon a second opening (H) in the printing frame is uniform from end to end. The actual variation of the illumination over the aperture areas was

[4] F. C. Toy, "The Measurement of Photographic Densities," Phot. J. **49**, 164 (1925).

[5] W. B. Ferguson, "The Standardization of Measurements of Photographic Density," Phot. J. **50**, 294 (1926).

definitely determined by means of a calibrated barrier-layer type photoelectric cell, the sensitive surface of which was masked except for a narrow slit. This cell was placed into the printing frame with the slit in the plane normally occupied by the printing material and oriented vertically. Next the cell was moved from point to point horizontally along the length of the aperture and

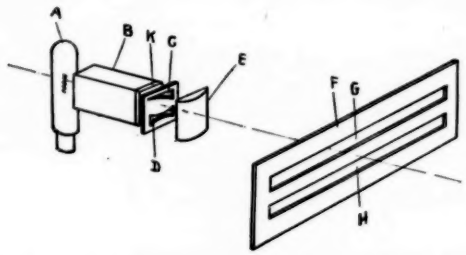


FIG. 2. Optical system for determination of contact printing density.

for each point of the aperture the light intensity was determined from the photoelectric current generated by the cell.

In order to understand the operation of this device, the printing frame, with its two illuminated rectangular apertures should be considered alone, since the remainder of the instrument constitutes simply a means towards obtaining an end. The aperture over which the illumination varies from end to end in a predetermined manner will serve to produce a calibrating exposure upon any printing material placed in the frame. The ratio of the illuminations at the extreme ends of this aperture was of the order of 100:1 which corresponds to a density range of 2.0. The uniformly illuminated aperture was arranged to have the same illumination as the bright end of the calibrating aperture. If now a negative is placed into the frame so as to cover only the uniformly illuminated aperture and if then a sheet of some printing material is inserted in the usual manner so as to cover both apertures, a print will be obtained upon which will appear the calibrating exposure as well as the exposure through the negative. If the negative has a printing density of between zero and two it will be possible to find a point on the calibrating exposure which has the same image density as the portion of the print exposed through the negative. From the photoelectric calibration the light intensity at this point can be determined; the intensity incident upon the uniformly illuminated aperture is already known; it remains only to insert these values into the Hurter and Driffield equation to obtain the density of the negative as recorded by the contact print.

In the above device all the rules of photographic photometry are observed. Calibrating and unknown exposures are made by means of the same light source on adjacent portions of the same material, and times of exposure for the two are identical. Prints were brush-developed, thus insuring as nearly as possible identical treatment for both exposures.

By this method a series of photographic images in positive film and supersensitive panchromatic film were measured, prints being made on three materials, (1) positive film, (2) matte P.M.C., and (3) glossy P.M.C.

In order that the results obtained might be compared directly with those of Toy and Ferguson, the next step in experimentation digressed from the primary purpose of this section. The photographic images were next evaluated by means of a *pot opal densitometer*, the essentials of which are shown in the diagram of Fig. 3 and whose optical system will be described here at the risk of possible redundancy, since it must be stressed that slight deviations from the specifications of optical systems used in most densitometers will lead to discrepancies in the results. Light from a source (*A*) was directed by means of the lens (*B*) upon a pot opal disk (*C*). A black mask (*D*) was painted on the second surface of the pot opal, permitting light to pass only through an aperture of approximately one quarter inch diameter. The photographic image (*E*) was placed over this aperture and the transmitted light measured by means of a calibrated barrier-layer type photoelectric cell (*G*) which was also masked to an aperture of one-quarter inch. With regard to the use of this cell, more will be said later.

A typical curve obtained by plotting the pot opal densitometer results against those of the contact printing densitometer is shown in Fig. 4, which represents the determination for supersensitive panchromatic film as printed on glossy P.M.C. paper. This curve, which is one of the

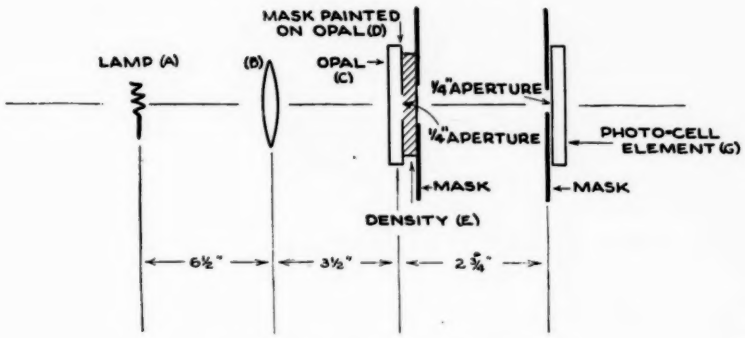


FIG. 3. Optical system of opal densitometer.

poorer ones so far as the accuracy of individual points is concerned but which otherwise is quite representative of the group, indicates a very nearly linear relation existing between the two sets of readings and, in general, the following equation may be used to link the pot opal readings with those of the contact printing device:

$$D_{pr}=AD_{po}+B$$

where A and B are constants representing the slope and the intercept on the D_{pr} axis, respectively. On this basis all the results may be tabulated as in Table I. Using the average values of A and B, one arrives finally at the equation:

$$D_{pr}=0.98D_{po}+0.04.$$

This result differs somewhat from those of Toy and Ferguson in that in their results the slope is less and the intercept does not appear. Since their findings are not duplicated, it was decided to repeat the experiment along the same lines as made by Toy with one slight difference: where Toy took the density of the film base as zero density, in the present work zero density represents the entire removal of the film. The results of this second determination are shown in Table II, from which we arrive at the following equation:

$$D_{pr}=0.97D_{po}+0.04.$$

These data indicate a slightly lower slope than the first set, but the intercept remains. Neither set of data shows any significant difference between results obtained with the different materials, which fact is in agreement with Toy's results.

Having made the above comparison with previous work on contact printing density, we shall now return to the original purpose of this section. At the moment we are primarily interested in integrating-sphere readings and not pot opal densitometer readings. The photographic images used in the contact printing experiment were next evaluated by means of an integrating sphere densitometer, shown in Fig. 5, which is self-explanatory, except, perhaps, that the function of the screen within the sphere is not immediately apparent. This point will be discussed later. Since, from the above results, it appears that no significant difference exists in the printing densities obtained with different printing materials and, since the results of the two determinations reported are so nearly alike, all the values obtained for the printing

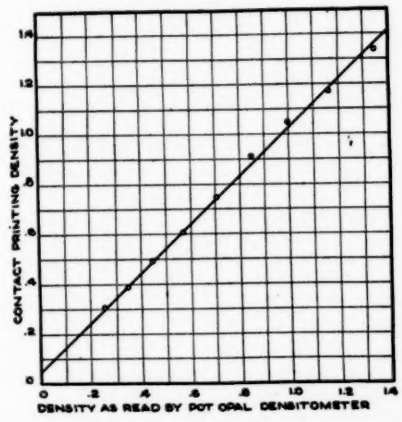


FIG. 4. Relation between readings of pot opal densitometer and contact printing density. Negative on S. S. Pan. Film. Print on glossy P.M.C.

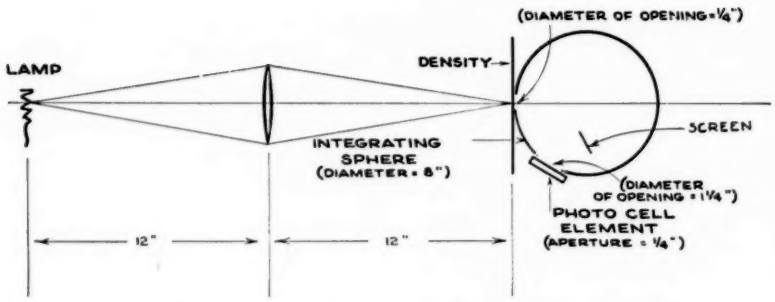


FIG. 5. Optical system of integrating sphere densitometer.

densities of the various photographic images have been averaged and in Figs. 6 and 7 are plotted against integrating-sphere determinations. In this manner, the following relations are arrived at:

$$\text{For positive film: } D_{pr}=1.03D+0.03.$$
$$\text{For S. S. pan. film: } D_{pr}=1.02D+0.03.$$

where D is integrating sphere density. In the next section it will be shown that the relation between D and D_{po} is not a linear one, especially at low densities. Hence, the equations obtained in this section cannot all be linear. From what follows in the remainder of this paper it will appear reasonable that a nonlinearity should exist in the relation between printing densities and integrating-sphere densities, especially at low values. Apparently the departure from linearity is too small to be detected unless a higher degree of accuracy is attained in determining printing density. Our results are not accurate enough either to confirm or to deny the existence of the exponential function found in the next section.

3. The Integrating Sphere and the Heavy Pot Opal Densitometer

From the discussion it appears that the integrating sphere has much to recommend it as a standard means of measuring photographic density. Unfortunately, however, the vast majority of densitometers in use today are not of this type, and to suggest that all these instruments should be replaced by integrating spheres would be entirely out of the question. At best, then, the integrating sphere can serve only as a primary standard.

In selecting a type of instrument to serve as a substandard, one is naturally led to the conclusion that the opal densitometer, with a heavy pot opal as diffuser, is the most logical, first, by the weight of sheer numbers, and second, because it is known to produce readings which approach what we have here chosen to call the true density of the photographic image.

If this instrument is to be considered as a substandard, it is necessary to determine more carefully the relations existing between its readings and density as evaluated by means of the integrating sphere.

The optical systems of the opal densitometer and the integrating-sphere densitometer have already been described. The measurements to be reported here and in the remainder of the paper need further explanation. It will be noted that all measurements were made by means of a calibrated barrier-layer type photoelectric cell. The glass window of this cell was removed so that

TABLE I.

Image Material	Print Material	A	B
Positive Film	Positive Film	0.98	0.03
Positive Film	Matte P.M.C.	.98	.04
Positive Film	Glossy P.M.C.	.98	.06
S. S. Pan. Film	Positive Film	.98	.04
S. S. Pan. Film	Matte P.M.C.	.97	.04
S. S. Pan. Film	Glossy P.M.C.	.98	.05

TABLE II.

Image Material	Print Material	A	B
Positive Film	Positive Film	0.99	0.04
Positive Film	Matte P.M.C.	.96	.05
Positive Film	Glossy P.M.C.	.97	.04
S. S. Pan. Film	Positive Film	.96	.05
S. S. Pan. Film	Matte P.M.C.	.96	.02
S. S. Pan. Film	Glossy P.M.C.	.95	.05

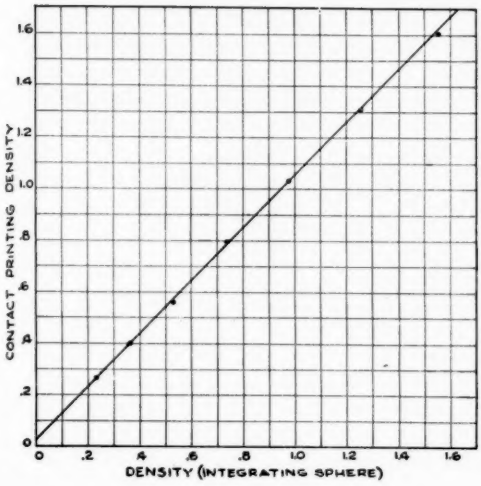


FIG. 6. Relation between readings of integrating sphere densitometer and contact printing density for photographic images on positive film.

radiation fell directly onto the sensitive surface. The aperture was kept constant at approximately one-quarter inch, for all measurements and illumination over this area was kept as uniform as possible so that no errors in calibration should be encountered.

In measuring the incident light and also the lower densities by means of the sphere, the incident beam is not at all or only slightly scattered and as a result an intense bright spot appears on the wall of the sphere. When images of higher densities are measured, more light is scattered from the narrow beam and the contrast between the spot and the surrounding areas decreases. In order to avoid an error due to a large change, resulting from this effect, in the angle of incidence of the light falling on the photoelectric cell, a white screen was inserted into the sphere between the photoelectric cell and that portion of the wall which was spot-illuminated. If this precaution is observed, the direction of illumination of the cell remains practically unchanged for all measurements.

The actual calibration of the cell was accomplished by means of the inverse square law: that is, the cell, together with an approximate point source of light, was set up on an optical bench, the distance between the source and the cell was varied, and readings of the photoelectric current were obtained by means of a galvanometer. The cell was recalibrated frequently, but the variations over a long period of time were found to be negligible. The sensitivity of the galvanometer used in connection with the cell (0.0001 microampere per millimeter) was such that it was possible to work at sufficiently low light intensities so that no fatigue effects were noticed in the photoelectric currents generated. This is the principal reason why only densities in the range 0–2.0 are discussed in this paper. A density reading of 2.0 corresponded to a deflection of about 5 mm on the galvanometer scale.

To guard further against errors resulting from cell characteristics, all readings were made with very nearly the same initial light intensities. As to the reproducibility of results, it may be pointed out that it was possible to repeat, after several months, a set of readings with an accuracy of ± 0.01 in density.

In referring to results obtained with an opal densitometer, it is essential that the completeness

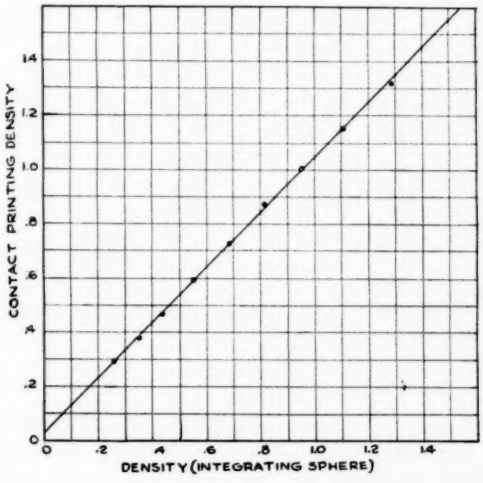


FIG. 7. Relation between readings of integrating sphere densitometer and contact printing density for photographic images on S. S. pan. film.

of diffusion provided by the particular opal used should be indicated. The reason for this will appear later in a discussion of the mechanism of the opal densitometer. The usual method of specifying opals is by means of a graph showing

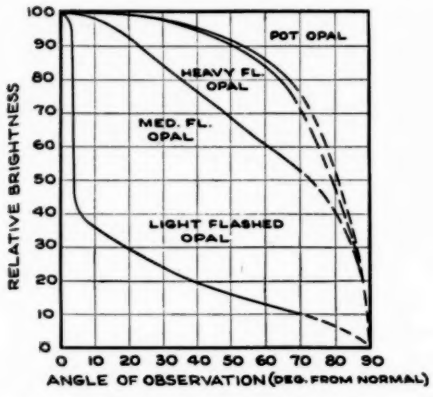

POT OPAL
HEAVY FL. OPAL
MED. FL. OPAL
LIGHT FLASHED OPAL
RELATIVE BRIGHTNESS
ANGLE OF OBSERVATION (DEG. FROM NORMAL)
0 10 20 30 40 50 60 70 80 90
0 10 20 30 40 50 60 70 80 90 100

FIG. 8. Light distribution curves for opal glasses. Diffusion efficiencies: Pot opal, 84.5; heavy flashed opal, 82.7; medium flashed opal, 71.0; light flashed opal, 22.0.

the brightness of the illuminated opal as a function of the angle of observation. From this curve the diffusion efficiency,[6] which is the figure obtained by dividing the area under the brightness distribution curve for the sample into the area under the curve for a perfect diffuser, may be obtained. The curve for a perfect diffuser is, by definition, a straight line parallel to the axis of abscissae. The distribution curves and diffusion efficiencies of the pot opal used here, as well as for other opals which will be referred to later, are shown in Fig. 8.

In comparing the readings of various densitometers here, and later in this paper, we have measured photographic images in several emulsions and on several bases. Specifically, figures have been obtained for a series of images in three samples of positive film, three samples of supersensitive panchromatic film, a lantern-slide plate, and a panchromatic press plate. Thus, two extremes, as well as intermediate values, of grain size are represented. These photographic samples were sufficiently neutral in color so that color filters having widely different characteristics produced no difference in density readings when inserted into the light path.

In Fig. 9 the readings obtained by means of the opal densitometer, with heavy pot opal as a diffuser, are plotted against readings obtained by means of the integrating sphere. It is immediately obvious that the relation existing between the two sets of data is not a simple linear one, since there is a marked curvature at low values of density. It will be found, upon analysis, that the relation differs from a linear one by an amount which is exponentially related to the readings of the sphere densitometer. Mathematically, the equation relating the two sets of data is as follows:

$$D_{po}=1.045D-0.034\epsilon^{-2.5D}.$$

This equation holds with good accuracy for all the materials tested.

It is unfortunate that the relation connecting the readings of a pot opal densitometer and true density should be so complex. The significance of the factors involved in the equation will be discussed later.

4. Other Densitometric Systems

Having determined the relationship of the previous section, we shall, in what follows, refer all other readings to those of the pot opal densitometer. We do this for two reasons: (1) Recorded in this manner, the data will have the maximum practical value and, (2) if readings of the remaining densitometers which we wish to examine are referred to sphere readings, we will

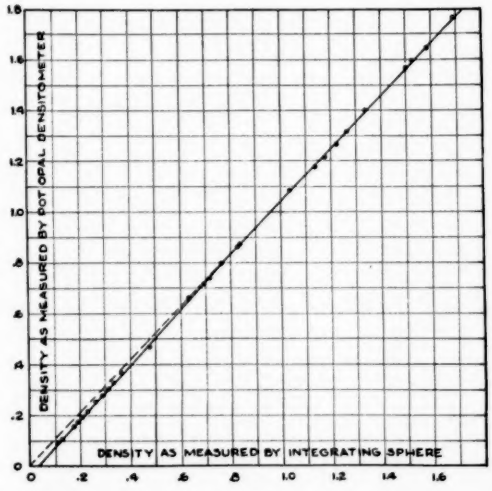

DENSITY AS MEASURED BY POT OPAL DENSITOMETER
DENSITY AS MEASURED BY INTEGRATING SPHERE
0 .2 .4 .6 .8 1.0 1.2 1.4 1.6
0 .2 .4 .6 .8 1.0 1.2 1.4 1.6 1.8

FIG. 9. Relation between readings of pot opal densitometer and integrating sphere densitometer.

[6] C. M. Tuttle and D. A. Young, "Illumination in Projection Printing of Motion Pictures," J. Soc. Mot. Pict. Eng. **19**, 842 (1932).

find an unwieldy exponential relation similar to the one just determined to apply in each case, whereas, if comparisons are made with pot opal determinations, a simple linear relation usually results.

The remaining instruments which we propose to examine are of two principal types: (1) those in which the incident light is supplied by some other form of opal glass, and (2) those in which the incident light is collimated and an attempt is made at collection of the total transmitted light by means other than the integrating sphere.

The opal instrument used in this section is identical with the one already completely specified, except for the diffuser used. Two examples of the second class are shown in Figs. 10 and 11. In the first the photographic image is laid directly against the surface of the light-sensitive element of a barrier-layer type photoelectric cell and illuminated by means of a collimated beam of light. This arrangement has been suggested by Hiatt and Tuttle[7] who also found that the results obtained therewith were in excellent agreement with visual measurements made on a pot opal densitometer. In the second example the image was placed against the polished end of a quartz rod of approximately one-quarter inch diameter and $1\frac{1}{2}$ inches in length, which serves to conduct the transmitted light from the image to the cell. This arrangement has been used in certain applications at these laboratories.

The following diffusers were tested in the opal densitometer: (1) heavy flashed opal, matte; (2) heavy flashed opal, polished; (3) medium flashed opal, matte; and (4) light flashed opal, polished. It was found that no appreciable difference appeared in opals having the same degree of flashing but different surfaces (matte or polished) so that, in what follows, no reference will be made to the surface of the opals. The light-distribution curves for the opals have already been indicated.

In this section the data obtained by means of the printing densitometer and the integrating sphere will also be recalculated, this time taking the readings of the pot opal instrument as the *independent* variable.

[7] B. C. Hiatt and C. M. Tuttle, "Note on the Measurement of Photographic Densities with a Barrier-Type Photocell," J. Soc. Mot. Pict. Eng. **26**, 195 (1936).

The procedure of plotting a set of curves from which equations are obtained has been followed again. It will not be necessary to present the curves herewith; only the equations will be given. With regard to the accuracy of the determinations, very few of the experimental points used in determining the curves are removed from the curves by an amount greater than 0.01 in either direction. The equations given below have

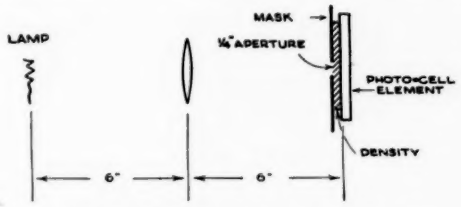


FIG. 10. Simple densitometer using barrier-type photoelectric cell.

been used to predict readings of one densitometer from those of another with an extreme error of ± 0.01 in density. (This does not hold for the contact printing densities where the accuracy is considerably less, but one should still be able to predict values with an error of not more than ± 0.02 in density.)

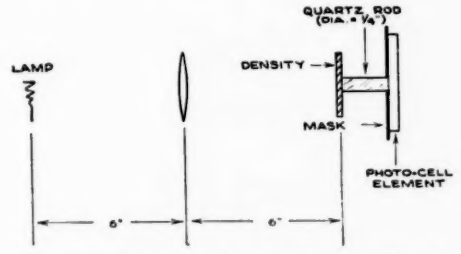


FIG. 11. Densitometer using quartz rod and barrier-type photoelectric cell.

Upon analysis it was found that all the curves could be made to fall naturally into one of three classes. The first of these is represented by a linear equation:

$$D_x = AD_{po} \pm B,$$

the second by a linear equation plus an exponential decay function:

$$D_x = AD_{po} + B\epsilon^{-ED_{po}},$$

and the third by an exponential function plus a

TABLE III. $D_x = AD_{po} + B$.

MATERIAL	HEAVY FLASHED OPAL		MEDIUM FLASHED OPAL		SPHERE—BASE TOWARDS SPHERE		PHOTOCELL CONTACT		PHOTOCELL QUARTZ ROD	
	A	*B*	*A*	*B*	*A*	*B*	*A*	*B*	*A*	*B*
Positive Film	1.0	0.012	1.027	0.017	0.957	0.032	1.0	0.015	1.035	0.027
S. S. Pan. Film	1.0	.012	1.036	.017	.957	.032	1.0	.010	1.025	.017
Lantern Slide Plate	1.005	.014	1.030	.023					1.025	.026
Pan Press Plate	1.005	.014	1.027	.020					1.025	.026

constant:

$$D_x = AD_{po}{}^E + B.$$

In the Tables III, IV, and V, the data are classified according to the above equations and the constants are given for each of the densitometers when used with the various materials tested.

5. INTERPRETATION OF RESULTS

Before attempting to interpret the results obtained so far, it will be well to consider the mechanism of operation of the opal densitometer, since in so doing we shall encounter most of the factors which produce the differences in results obtained with various instruments.

In the opal densitometer a beam of light incident upon the opal is diffused thereby. The diffused light is then incident upon the photographic image and its associated base. Here a portion of the light will be reflected by image and base, and the remainder will be either absorbed or transmitted. Now all the light which is reflected from the film, say, is not lost, for upon striking the opal, which has a high reflection coefficient, a portion of it is rereflected to be once more incident upon the film. Now, if, as is usual, the light incident upon the film is measured by removing the film from the opal and determining the unmodified brightness of the opal, it is apparent that a small amount of light, due to interreflection, which is incident upon the film, remains unmeasured so that an error is introduced into the readings.

TABLE IV. $D_x = AD_{po} + B\epsilon^{-ED_{po}}$.

MATERIAL	SPHERE—EMULSION TOWARDS SPHERE		
	A	*B*	*E*
Positive Film	0.957	0.030	2.5
S. S. Pan. Film	.957	.030	2.5
Lantern Slide Plate	.957	.030	2.5
Pan Press Plate	.957	.030	2.5

TABLE V. $D_x = AD_{po}{}^E + B$.

MATERIAL	LIGHT FLASHED OPAL		
	A	B	E
Positive Film	1.23	0.03	0.96
S. S. Pan. Film	1.23	.02	.96
Lantern Slide Plate	1.23	.02	.96
Pan Press Plate	1.23	.02	.96

Next, in order to understand the function of the opal, consider for a moment a single beam of light to be normally incident upon the photographic image. The light transmitted by the image will be made up of a principal beam having the same direction as the incident one, and a number of scattered rays, which leave the image at various angles to the incident beam. If a light-collecting device of small aperture, such as the eye, is placed in the path of the principal beam at some distance from the image, it will fail to collect the scattered rays and will fall short in appraising the quantity of light leaving the image. Now suppose that a second beam of the same intensity as the first is made to be incident upon the image at the same point as the first but making an angle with it. Then the principal ray from the second beam will not enter the eye, the position of which has remained the same. However, the direction of this second beam may be such that one of the rays scattered from its original direction *will* enter the eye. The scattered ray thus measured will correspond exactly with one of the rays which was scattered *away* from

the first beam and which the eye *failed* to measure. Finally, suppose that an infinite number of rays of equal magnitude are incident on the film in all directions. Then for each ray scattered from the perpendicular beam *away* from the eye, a similar one will be scattered from one of the other beams *into* the eye and the results should be the same as if only one beam were used and the eye were able to collect all the rays scattered therefrom. It is the function of the opal to provide the infinite number of rays, each in a different direction, all of equal magnitude. If the diffuser is not a perfect one, the result may be that not all of the scattered light is measured, and since we know of no perfect diffuser, it must be concluded that the opal densitometer may fail to measure a small portion of the scattered light.

It appears, then, that the opal densitometer has two principal inherent defects, the first arising from the interreflection effect, and the second from the failure of the opal to diffuse perfectly.

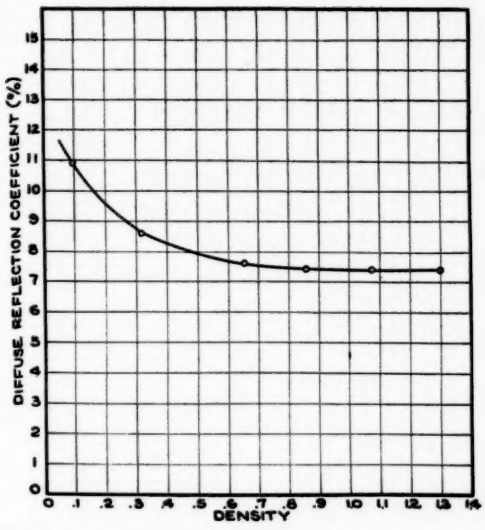


FIG. 12. Relation of reflection coefficient to density of photographic image for images in positive film.

Now the scattering of light by the photographic image increases as the density of the image increases. At zero density there is theoretically no scattering. Failure of the opal to approach perfection will therefore not affect the low density readings, but at high densities it will result in a reading which is too high. Hence, if true density were to be plotted against opal readings, we should expect the curve to have a slope less than unity. If, on the other hand, readings obtained with a very poor diffuser were to be plotted against similar readings obtained with a good diffuser, such as pot opal, the curve would be expected to have a slope greater than unity. That this is actually the case appears from the values for (A), the slope, obtained when readings of various instruments are plotted against pot opal readings, and given in Table VI. Thus, we see that the slope is indicative of the completeness with which scattered light is measured.

The interreflection effect appears in a different manner. It operates at very low values of density since, although there may be no silver deposit, the base itself reflects more than 10 percent of the light incident thereon. As a matter of fact the clear base is a better reflector than is a film bearing a photographic image, as is demonstrated by the curve of Fig. 12, where diffuse reflection coefficients obtained by means of an integrating sphere are plotted against density. On this basis we would expect, in plotting true density against opal densities, to find a positive intercept on the ordinate axis. Such an intercept appears in the equations given above where it is represented by the constant (B). If values of (B), together with the diffuse reflection coefficients of the opals are tabulated for various densitometers, the results given in Table VII are obtained. Hence the intercept is, in some measure at least, related to the magnitude of the interreflection effect.

It can be shown, in the case of an opal densitometer, that if:

I_0 = intensity of light incident on film,
I = intensity of light transmitted by film,
R_1 = reflection coefficient of film,
R_2 = reflection coefficient of diffuser,
and T = transmission coefficient of film,

then:

$$I = I_0 T(1 - R_1 + R_1R_2 - R_1^2R_2 + R_1^2R_2^2 - \cdots).$$

If we assume that $R_1 = 0.10$ and $R_2 = 0.70$, figures which are very near the actual ones when clear

base is measured on a pot opal densitometer, we arrive at

$$I=0.968I_0T.$$

In the case of the sphere, using the same symbols as above:

$$I=I_1T(1-R_1) \text{ very nearly,}$$

and, substituting the value assumed for R_1, we obtain:

$$I=0.9I_0T.$$

The interreflection effect is responsible for a difference in density given by $\log_{10} 0.968/.9$, the value of which will be found to be 0.032. This happens to be very nearly the intercept obtained in plotting integrating-sphere readings against pot opal readings.

If the reflection coefficient of the film were constant, the interreflection effect would appear simply as a subtracted constant increment in density, which in the case of the pot opal would be 0.032. However, we have shown above that for values of density up to one, the reflection coefficient of the film decreases as density increases; as a result, the subtracted increment decreases with density and we find that the relation between integrating-sphere readings and pot opal readings is not a linear one, especially at low values of density. The actual departure from linearity found experimentally is somewhat greater than would be predicted from the change in reflection coefficient alone, but the effect is probably accentuated due to the fact that the reflection from the film at high densities is very diffuse while at low densities it is quite specular.

TABLE VI.

DENSITOMETER	DIFFUSION EFFICIENCY	SLOPE
Int. Sphere	100	0.957
Heavy Fl. Opal	82.7	1.0
Medium Fl. Opal	71.0	1.027
Light Fl. Opal	22.0	1.23

The results obtained by means of the two densitometers of the second type can be analyzed on the above basis. In this manner we arrive at the conclusion that in the case where the photographic image is placed directly against the sensitive cell surface, the scattered light is measured with about the same degree of completeness as by means of the pot opal densitometer. That it is not measured as effectively as it would be by the integrating sphere may be explained by saying that the photoelectric element does not accept light at large angles of incidence as readily as it does the normal rays.[8] Where the quartz rod is inserted between the cell element and the image, still less of the scattered light is measured, perhaps because even less of the scattered light is accepted at the end of the rod, perhaps because some of the scattered light is lost as a consequence of its longer path through the cylinder. In both of these densitometers there is an interreflection effect which is less than in the case of the pot opal.

One would expect that in contact printing the interreflection effect would be nearly as great as in the case of the pot opal densitometer since printing materials are highly reflecting, but if the intercept is taken as a measure of interreflection, one must conclude that interreflection is less in contact printing than in the integrating sphere. Since this is obviously impossible, it becomes apparent that some other fact has entered to increase the apparent density of the base in the case of the contact print. Perhaps the intercept appears because the printing media used above are sensitive into the ultraviolet where the transmission of the base decreases rapidly.

In the opal densitometer, as the flashing of the opal is decreased, a point is finally reached where the curves connecting pot opal readings and the thin flashed opal readings become definitely nonlinear. Silberstein and Tuttle[9] have sho w mathematically that the relation connecting the

TABLE VII.

DENSITOMETER	REFLECTION COEFFICIENT	INTERCEPT
Integ. Sphere	negligible	0.030
Light Fl. Opal	15	.020
Medium Fl. Opal	22	.017
Heavy Fl. Opal	34	.012

[8] "Report of the Committee on Photoelectric Portable Photometers," Illum. Eng. Soc. September 3, 1936.

[9] L. Silberstein and C. M. Tuttle, "The Relation Between the Specular and the Diffuse Photographic Densities," J. Opt. Soc. Am. **14**, 365 (1927).

diffuse and specular densities is an exponential one and therefore it is quite logical that curvature would begin to appear with thin flashed opals. Bloch and Renwick[10] also found the exponential relation to hold experimentally, as did Tuttle.[11] In the present study, however, we find that, where readings of thin opals are to be connected with pot opal readings, the relation is not a purely exponential one, but that a constant must be added to take care of the interreflection effect.

It is surprising to find, in all the above results, that practically no difference exists in the slope constant obtained for photographic materials having widely different grain sizes, this condition extending even to the very thin flashed opal curves.

TABLE VIII.

NORMAL	REVERSED	DIFFERENCE
Density as Read by Integrating Sphere		
1.785	1.819	0.034
1.524	1.558	.034
1.006	1.041	.035
.761	.790	.029
.561	.586	.025
.390	.410	.020
.250	.265	.015
.172	.184	.012
Density as Read by Pot Opal Densitometer		
1.657	1.685	0.028
1.209	1.238	.029
.815	.841	.026
.502	.523	.021
.306	.315	.009
.181	.189	.008
.117	.118	.001

6. EFFECT OF REVERSING ORIENTATION OF IMAGE WITH RESPECT TO BASE IN MEASURING DENSITY

So far nothing has been specified regarding the orientation of the film in any of the above measurements, that is, whether the base side or the emulsion side faced the light source. When measurements are made on an opal densitometer, the emulsion is usually placed in contact with the opal; in measurements with the integrating sphere it is usual, in order that all the scattered light shall be collected, to place the emulsion towards the sphere; in contact printing the emulsion should, obviously, be placed against the printing medium. These are the orientations which were used in the experiments reported thus far.

If, in the case of the sphere and the opal densitometers, the film is reversed from the above orientation, the effect is to increase the density, by an amount which is indicated in Table VIII.

The increase in density is not due simply to the fact that the image has been removed by several thousandths of an inch from the opal or the sphere, as the case may be, since if the image is removed a like distance without reversing the film—a case which may be tried experimentally with an annular ring shim of thickness equal to that of the base—no change in density occurs.

When the effect was first discovered with the integrating sphere, it was thought that opposite sides of the film might have quite different reflection coefficients, but later it was found that the specular density of the film remained the same regardless of orientation. Since, in a specular determination scattered light is not measured, it must be concluded that the difference in density upon reversal of the film is intimately connected with scatter, and not with reflection. Measurements of the diffuse reflection coefficients of opposite sides of film samples bear out this theory, since no difference appears.

In the case of the pot opal densitometer, when the film is reversed so that the base is towards the opal, some light from the opal traverses the base at very large angles, and is then scattered by the emulsion into the measuring device. Somewhat similarly, in the case of the sphere, when the base is towards the sphere, light scattered by the emulsion travels through the base at large angles, to be finally collected by the sphere. In both cases the determination of higher densities is partially dependent upon the collection and measurement of light which travels through the base at large angles. With the film in the normal position this is not true, since in both instruments the measured light travels

[10] O. Bloch and F. F. Renwick, "The Opacity of Diffusing Media," Phot. J. **40**, 49 (1916).

[11] C. M. Tuttle, "Densitometry and Photographic Printing. Illumination of the Negative and Its Effect Upon Density," J. Opt. Soc. Am. **24**, 272 (1934).

through the base perpendicularly. Now it is quite certain that some of the light which is scattered within the emulsion travels through the base at such large angles that it strikes the surface at angles greater than the critical angle and is thus not emitted. Again, since the base itself is not perfectly transmitting, a thickness

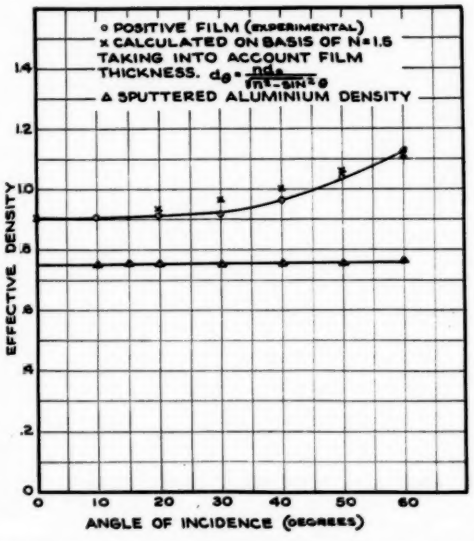


FIG. 13. Relation between effective density and angle of incidence of measuring beam.

correction must be applied to the rays traveling through it at other than normal angles (see Section 1). Thus we see that there is some reason to expect just such a change in density as actually takes place.

In this connection it is interesting to note that if densities determined by means of the sphere with the film reversed are compared with corresponding densities as determined by means of the pot opal with the film in the normal orientation, a linear relation exists between the two. The nonlinearity introduced by reversing the film, in the case of the sphere determination, almost exactly counterbalances the nonlinearity due to interreflection in the opal densitometer. Under the conditions specified here, the relation between pot opal readings and integrating sphere readings is very nearly as follows for both positive and supersensitive panchromatic emulsions:

$$D=0.957D_{po}+0.032.$$

7. Modification of the Distribution Curve of an Opal by a Density in Contact Therewith

The fact that the shape of the distribution curve of an opal glass is different from the shape of a corresponding curve for the same opal in contact with a photographic density has been reported from time to time and Bull and Cartwright published curves showing the effect.[12] They concluded that if the opal were a perfect diffuser, then the addition of a photographic density, a very imperfect diffuser, could not possibly change the distribution curve, and that the reason the effect occurred was because the opal was an imperfect diffuser. This is not the case, as has been pointed out by Selwyn and Pitt.[13] It is only necessary to apply Beer's law to the rays leaving the opal at an oblique angle and passing through the density obliquely to arrive at the curves obtained by Bull and Cartwright, i.e., rays passing through the film perpendicularly travel a much shorter distance through the density than do the oblique rays and are hence less attenuated. Now according to Beer's law, the density should be directly proportional to the length of path, for a nonscattering material. We have stated previously that the length of path, taking into account the refractive index, n, of the emulsion is given by:

$$l=nt/(n^2-\sin^2\theta)^{\frac{1}{2}},$$

where t is the thickness of the emulsion and θ is the angle from the normal. Hence, the effective density at any angle should be given by:

$$d_\theta=d_0n/(n^2-\sin^2\theta)^{\frac{1}{2}}.$$

Densities calculated from the above formula are in good agreement with the experimental values, as is shown by the curves of Fig. 13, and we must conclude therefore that the phenomenon is not connected with the light-scattering properties of the photographic image, but is almost entirely a thickness effect.

It should be noted that if a small aperture is used to limit the cone of light measured in an

[12] A. J. Bull and H. M. Cartwright, "An Evaluation of the Light Scattered by Photographic Densities," Phot. J. **49**, 177 (1925).

[13] E. W. H. Selwyn and F. H. G. Pitt, paper to be published.

opal densitometer, and if this aperture is at an appreciable distance from the photographic image, the modification of the distribution curve has no effect on the evaluation of density. This restriction in the size of the aperture is a requirement for the opal densitometer analogous to the collimation requirement in the case of the sphere instrument.

8. The Use of Thin Metallic Films as Standard Densities

It has been suggested that for the purposes of the calibration of densitometers, densities consisting of metallic surfaces sputtered onto glass bases be used as a standard. From observations made at these laboratories, it has been concluded that such metallic films have no value as standards, for the following reasons: (1) Metallic films do not behave as do photographic densities insofar as scattering is concerned; practically no light is scattered. (2) Incident light is modified almost entirely by pure reflection and not by absorption. For this reason light incident at an angle is no more attenuated in passage through the layer than is normally incident light. Experiments have shown that the effective density is constant even up to angles of incidence as large as 60° (see Fig. 13). (3) The metallic films are far from neutral in color.

The most serious of these objections is probably (2). Since the coefficient of reflection is high, the interreflection effect encountered when such a film is measured by means of a pot opal is enormous. It will be found therefore that pot opal densitometer readings are far below similar readings made on an integrating sphere densitometer, as is shown in Fig. 14 where sputtered aluminum films have been measured. On the other hand, if a very thin flashed opal is used, the density readings are higher than both integrating sphere and pot opal readings. This is because thin flashed opals specularly transmit the red but scatter the blue. The metallic films, on the other hand, absorb the red, so that measurements made in this manner indicate a high value of density. These facts should be sufficient to bar the metallic film from use as a standard density.

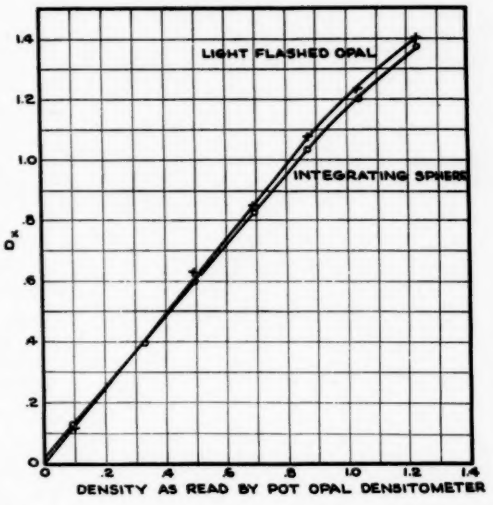


Fig. 14. Densities of thin metallic films as measured by various densitometers.

9. Conclusions

That standardization of densitometry throughout the photographic laboratories of the world is desirable cannot be questioned and the authors wish the experimental data presented in this paper to be considered as a contribution toward the solution of this problem.

As a primary standard instrument, the integrating sphere has two important arguments in its favor:

(1) Its measurements conform to the most logical interpretation of the Hurter and Driffield definition.

(2) The sphere and its illuminating optical system may be definitely specified—but rigid adherence to specification is not required since reasonably wide variations from the specifications do not influence the results.

The acceptance of the sphere as the primary standard instrument is made easier as a consequence of the fact that accurate calibration of substandard instruments of many varieties can be accomplished through the medium of a previously measured set of photographic images covering the desired range of densities. The calibration would, of course, consist in arbitrary scaling of an intensity-varying mechanism (which is an essential part of almost every practical densitometer) to give readings coincident

TABLE IX. *Density as Evaluated by Various Densitometers.*

Pot Opal	Pot Opal (Base to Opal)	Heavy Flashed Opal	Medium Flashed Opal	Light Flashed Opal	Sphere (Emulsion to Sphere)	Sphere (Base to Sphere)	Photocell (Image in Contact with Element)	Photocell (Quartz Rod)	Contact Printing
0.000	0.00	0.012	0.017	0.020	0.030	0.032	0.012	0.022	0.04
0.100	.107	.112	.120	.155	.119	.128	.112	.125	.14
0.200	.212	.212	.223	.282	.209	.223	.212	.229	.24
0.300	.316	.312	.326	.407	.302	.319	.312	.332	.33
0.400	.420	.412	.429	.530	.394	.415	.412	.434	.43
0.500	.522	.512	.532	.652	.487	.510	.512	.537	.53
0.600	.623	.612	.635	.773	.581	.606	.612	.641	.63
0.700	.724	.712	.738	.893	.675	.702	.712	.744	.73
0.800	.825	.812	.841	1.013	.769	.797	.812	.847	.82
0.900	.926	.912	.944	1.131	.864	.893	.912	.951	.92
1.000	1.027	1.012	1.047	1.250	.959	.989	1.012	1.054	1.02
1.100	1.128	1.112	1.150	1.368	1.054	1.084	1.112	1.157	1.12
1.200	1.228	1.212	1.253	1.485	1.149	1.180	1.212	1.260	1.22
1.300	1.328	1.312	1.356	1.602	1.245	1.275	1.312	1.374	1.31
1.400	1.429	1.412	1.459	1.719	1.340	1.371	1.412	1.477	1.41
1.500	1.529	1.512	1.562	1.835	1.436	1.467	1.512	1.580	1.51
1.600	1.629	1.612	1.665	1.951	1.531	1.563	1.612	1.684	1.61
1.700	1.730	1.712	1.768	2.067	1.626	1.659	1.712	1.787	1.71
1.800	1.830	1.812	1.871	2.183	1.722	1.754	1.812	1.890	1.80
1.900	1.930	1.912	1.974	2.298	1.818	1.850	1.912	1.994	1.90
2.000	2.030	2.012	2.077	2.413	1.914	1.946	2.012	2.097	2.00

with the predetermined values for the sample as determined on the integrating sphere.

To meet the requirements for a satisfactory secondary instrument, it is necessary only that the relation between apparent density as measured by the system and true density as measured by the sphere shall be the same for all photographic materials. This requirement is met by all of the opal densitometers as well as by the two semi-integrating types.

Table IX gives a comparison of the numerical values which would be obtained in densitometers having the different optical arrangements which have been discussed if such densitometers were to be calibrated on the basis of the transmitted to incident light intensity ratio.

A Refractometer for the Near Infrared

RUDOLF KINGSLAKE AND HILDA G. CONRADY
Institute of Applied Optics, University of Rochester, Rochester, N. Y.
(Received March 18, 1937)

A convenient apparatus is described for making refractive index measurements on a 60° prism, in the near infrared region. The instrument is essentially a double Wadsworth mirror monochromator, the first half being calibrated in wave-lengths, and the second half calibrated to read angular rotations of the prism table on which the unknown prism is mounted. The wave-length drum was calibrated by working backwards from a known prism (rocksalt or quartz) mounted on the second prism table. Measurements on eleven optical glasses are given, from 0.365 to 2.6μ.

WITH the recent advances in infrared sensitive photographic materials, the need is arising for lenses achromatized in the near infrared. To design such lenses, it is of course necessary to know the indices of the glasses it is proposed to use, up to wave-lengths of one or two microns, which are not usually given by the manufacturers of optical glass. Extrapolation[1] by means of the Hartmann formula has been used, but it cannot be carried very far.

The classical[2] measurements of refractive index in the infrared region have always been made with a prism spectrometer equipped with a 60° prism of the material under investigation, with either a mechanical linkage or Wadsworth mirror[3] arrangement to maintain the prism in minimum deviation at all wave-lengths.

In most cases, white light has been used, with some arrangement for measuring the mean wave-length of the narrow band of radiation passing through the spectrometer. The radiation is detected with a bolometer or thermopile. Langley[4] used in tandem with the prism spectrometer, a second spectrometer equipped with a concave grating, by means of which a series of known wave-lengths in successive orders could be sent into the prism. Paschen[5] made use of a plane grating in a similar manner. Rubens[6] employed a reflecting etalon at 45° to reflect light into the instrument, the spectrum being thus "channeled" with a succession of dark and bright bands falling at known wave-lengths. Lowry and Coode-Adams[7] used a similar method, but found it unsatisfactory beyond 2μ because of the high opacity of a silver layer at those wave-lengths. A transmission etalon can be used in a similar way, as suggested by Edser and Butler,[8] but both methods suffer from lack of precision on account of the great broadness of the bands. Mouton and Carvallo[9] made use of a quartz plate between crossed Nicols to give a channeled spectrum. With a sufficiently sensitive thermopile, known spectral lines[10] from gaseous discharge tubes may of course be used, as in ordinary visual refractometry.

However, since so many high precision measurements have been made on crystalline materials such as quartz or rocksalt, it is now possible to work backwards from the index of a prism composed of one of these materials, to the wave-length of the radiation corresponding to that index.

Thus if infrared light is known to be passing through such a prism of angle a, in minimum deviation, and if the angle of incidence is known to be i, the angle of refraction will be $a/2$, and hence the index must be given by $\sin i/\sin \frac{1}{2}a$.

From a table of given indices, plotted on a large-scale graph, the wave-length of the radiation may be at once inferred. In this way it is unnecessary to use gratings, and sufficient intensity is obtainable to operate a relatively crude and insensitive thermopile.

[1] Lee, J. Sci. Inst. **9**, 277 (1932).
[2] For a survey of the subject, see Baly, *Spectroscopy* 1, p. 216; Lecomte, *Le Spectre Infrarouge*, p. 97; and Schaefer and Matossi, *Das Ultrarote Spektrum*, p. 40.
[3] Wadsworth, Phil. Mag. [5] **38**, 337 (1894); also Gibbs and Collins, Astrophys. J. **56**, 483 (1922).
[4] Langley, Ann. d. Physik [3] **22**, 598 (1884).
[5] Paschen, Ann. d. Physik [3] **53**, 301 (1894).
[6] Rubens, Ann. d. Physik [3] **45**, 238 (1892).
[7] Lowry and Coode-Adams, Phil. Trans. Roy. Soc. **226**, 391 (1927).
[8] Edser and Butler, Proc. Phys. Soc. **16**, 207 (1898).
[9] Carvallo, J. de phys. **2**, 27 (1893).
[10] Paschen, Ann. d. Physik [4] **35**, 1005 (1911).

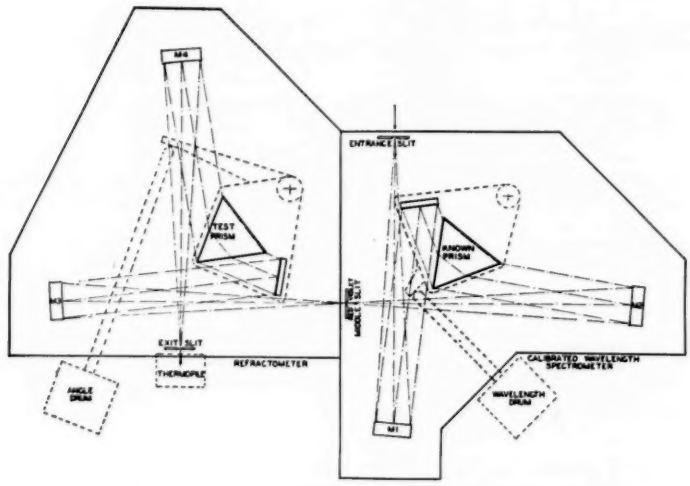


FIG. 1. General arrangement of the infrared refractometer.

AN INFRARED REFRACTOMETER

In the instrument recently completed at the University of Rochester, two Wadsworth-type prism monochromators are mounted in tandem to constitute a regular double monochromator. The first contains a fixed dense flint glass prism, while the second can be used with any known or unknown prism. The drum of the first monochromator containing the flint glass prism has been calibrated directly in wave-lengths, by a method explained below, while that of the second is marked to show the angular position of the turntable in minutes of arc. All lenses have been replaced by decentered parabolic concave mirrors to avoid chromatic aberration, the mirrors having a diameter of 38 mm and a focal length of 305 mm, thus giving a relative aperture of f: 8. A diagram of the system is shown in Fig. 1. White light from a 250-watt projection lamp, connected in series with a ballast lamp to maintain constant brightness, is focused on the first or entrance slit by means of a tilted concave mirror (not shown).

The dense flint prism ($n=1.71$) in this half of the instrument forms a continuous spectrum on the second slit, from which a narrow band of wave-lengths is selected and allowed to pass on through the second half of the system. The width of this wave-length band is of course finite and varies with the dispersion of the glass prism, the drum for rotating the prism table being marked with the wave-length of the centroid of the transmitted band. As long as the slits are narrow, no calibration error is likely to arise due to the nonlinear spectral energy distribution of the source, but with wide slits, the calibration may be somewhat in error from this cause except around the peak of the energy distribution curve. The "crowding" of the red end of the prismatic spectrum would also cause a calibration error when wide slits are used, but if the instrument is calibrated and used with slits of constant width, this effect will be automatically compensated. The slit width actually employed was 0.1 mm for all three slits, an Eppley copper-constantan thermopile and a Leeds and Northrup moving coil galvanometer being employed to detect the radiation emerging from the third slit.

The narrow band of "monochromatic" light given by the first instrument then passed through the unknown prism in the second half of the system. The unknown prism was first set at minimum deviation by visual observation for the yellow neon line at $\lambda=0.5852\mu$, this line being chosen as it is so easy to obtain from a cheap neon lamp run off the 120-volt supply. The table was turned to the point at which this line fell centrally on the exit slit. The thermopile was then substituted for the eyepiece, and the first monochromator drum was set at the desired infrared wave-length. The second prism table was then

turned slowly by its drum until the thermopile recorded a maximum of transmitted radiation. Since a Wadsworth mirror was being used, the light which finally found its way out of the third mirror had necessarily passed through the unknown prism in minimum deviation.

Thus the angle through which the second prism table had to be turned to go from the yellow neon line to the given infrared line, was equal to the change of angle of incidence of the prism from one wave-length to the other, the angle of refraction being always equal to the half-angle of the prism. If now the same unknown prism was removed from the apparatus and mounted on an ordinary laboratory spectrometer supplied with neon light, the angle of minimum deviation d could be measured for the yellow line, and the angle of incidence calculated by $i=\frac{1}{2}(d+a)$ where a is the prism angle. The angle of incidence, ϕ, for the infrared line could then be calculated by simple subtraction and the infrared index could be computed by $\sin\phi/\sin\frac{1}{2}a$. In practice, three or four visual lines were measured to act as checks on the measurement of di, which is the difference between the angle of incidence of the prism for neon yellow light and for any other specified wave-length, since the angle of minimum deviation of the prism could be also measured directly for those visible wave-lengths on an ordinary spectrometer.

For these checks, either discharge tubes giving the actual required lines could be used, or a white source could be employed, the calibrated monochromator being set to give a narrow wave-length band centered about each of the desired wave-lengths in turn.

Calibration of the Monochromator

The drum on the second half of the instrument was marked at each minute of arc rotation of the table by mounting a temporary plane mirror on the table, and obtaining an autocollimated image from the mirror in the telescope of an ordinary spectrometer placed close up to the apparatus. The dimensions of the apparatus were such that each minute of arc was represented by about 5 mm on the drum scale, the whole range of movement being 300 minutes (5°).

The first half of the apparatus, which is an ordinary monochromator, was calibrated in the visible region by making settings on a number of lines of known wave-length from the spectra of mercury, neon, helium, and argon. A large graph was then constructed connecting the wave-length with the reading on a micrometer head attached for this purpose to the wave-length drum. For the infrared calibration, the whole double instrument was used, a prism of some known crystalline material being mounted on the second turntable, in the position to be occupied eventually by the unknown prism. The prism table carrying this prism was then set at a series of positions corresponding to the computed angles of incidence for which index-wave-length data are available, and at each setting the wave-length drum of the monochromator part was turned until the thermopile showed a peak in the intensity of the emerging radiation. These points were then added to the calibration graph, and formed a continuation of the curve already obtained for the visible region.

Calibration was effected by means of a rocksalt and a quartz prism:

(a) *Rocksalt prism:* An excellent prism was supplied by the Bausch and Lomb Company, in which the surfaces were of good figure, but slightly convex. This was easily compensated by a trifling change in the focus of the concave mirrors, the spectral lines so obtained being practically as good as with a quartz or glass prism.

Apart from the physical difficulty of handling it, a rocksalt prism suffers from a high temperature coefficient, which according to Micheli[11] and Langley[12] has the values shown in Table I, in units of the fifth decimal of refractive index per degree centigrade.

Table I. *Temperature coefficient of rocksalt.*

λ	dn	λ	dn	λ	dn
0.361μ	−3.194	0.589	−3.622	1.123	−3.88
.441	−3.425	.643	−3.636	1.445	−3.88
.467	−3.454	.760	−3.73	1.950	−3.85
.480	−3.468	.833	−3.82	2.7	−3.84
.508	−3.517				

[11] Micheli, Ann. d. Physik **7**, 772 (1902).
[12] Langley, Ann. Smithson. Ast. Obs. **1**, 101 (1902).

TABLE II. *Refractive indices of optical glasses.*

Glass No. / Maker / Type No.	1 B-L 70	2 B-L 20–2	3 PM 5951	4 B-L 160–1	5 B-L 110–1	6 PM 6581	7 B-L 120–1	8 PM 4887	9 B-L 120–2	10 Schott F2	11 B-L 220–3
Name	Hard Crown	Boro-silicate Crown	Extra Light Flint	Tele-scope Flint	Dense Barium Crown	Light Flint	Dense Barium Crown	Light Barium Flint	Dense Barium Crown	Dense Flint	Extra Dense Flint
N_D	1.51166	1.51666	1.52390	1.52970	1.57136	1.57348	1.61173	1.60535	1.61703	1.62082	1.71968
N_F-N_C	.00845	.00802	.01039	.01028	.00995	.01338	.01037	.01381	.01141	.01707	.02454
V	60.6	64.4	50.4	51.5	57.4	42.9	59.0	43.8	54.1	36.4	29.4
Wave-length:											
0.3650μ	1.53245	(?)	1.55052	1.55555	1.59610	1.60880	1.63738	1.64171	1.64570	1.66714	(absorbed)
h .4046	1.52597	1.53013	1.54192	1.54736	1.58837	1.59718	1.62935	1.62975	1.63666	1.65163	1.76506
G' .4340	1.52239	1.52676	1.53728	1.54286	1.58407	1.59098	1.62491	1.62338	1.63164	1.64346	1.75270
g .4358	1.52217	1.52657	1.53703	1.54262	1.58382	1.59064	1.62467	1.62300	1.63138	1.64300	1.75205
F .4861	1.51760	1.52227	1.53125	1.53695	1.57838	1.58300	1.61903	1.61516	1.62508	1.63307	1.73732
e .5461	1.51375	1.51864	1.52645	1.53224	1.57380	1.57678	1.61429	1.60870	1.61982	1.62503	1.72569
D .5893	1.51166	1.51666	1.52390	1.52970	1.57136	1.57348	1.61173	1.60535	1.61703	1.62082	1.71968
C .6563	1.50915	1.51425	1.52086	1.52667	1.56843	1.56962	1.60866	1.60135	1.61367	1.61600	1.71278
.60	1.51123	1.51623	1.52336	1.52913	1.57081	1.57280	1.61118	1.60456	1.61638	1.62005	1.71842
.65	1.50937	1.51447	1.52112	1.52688	1.56867	1.56997	1.60891	1.60163	1.61394	1.61632	1.71330
.70	1.50789	1.51302	1.51934	1.52507	1.56690	1.56764	1.60709	1.59930	1.61196	1.61358	1.70935
.75	1.50659	1.51179	1.51788	1.52357	1.56542	1.56583	1.60557	1.59740	1.61030	1.61126	1.70608
.80	1.50548	1.51070	1.51658	1.52230	1.56420	1.56422	1.60423	1.59581	1.60890	1.60936	1.70358
.85	1.50458	1.50978	1.51548	1.52117	1.56314	1.56290	1.60312	1.59445	1.60773	1.60780	1.70140
.90	1.50373	1.50893	1.51454	1.52015	1.56217	1.56172	1.60212	1.59324	1.60668	1.60635	1.69956
.95	1.50299	1.50818	1.51367	1.51927	1.56137	1.56066	1.60123	1.59222	1.60576	1.60517	1.69794
1.00	1.50229	1.50745	1.51289	1.51844	1.56063	1.55975	1.60039	1.59128	1.60493	1.60408	1.69651
1.1	1.50104	1.50613	1.51147	1.51689	1.55921	1.55812	1.59891	1.58966	1.60343	1.60220	1.69408
1.2	1.49985	1.50484	1.51021	1.51547	1.55795	1.55663	1.59748	1.58823	1.60205	1.60055	1.69199
1.3	1.49873	1.50359	1.50902	1.51413	1.55683	1.55531	1.59624	1.58699	1.60083	1.59911	1.69028
1.4	1.49768	1.50238	1.50784	1.51288	1.55573	1.55403	1.59500	1.58578	1.59970	1.59779	1.68866
1.5	1.49665	1.50114	1.50677	1.51158	1.55465	1.55282	1.59380	1.58464	1.59854	1.59652	1.68725
1.6	1.49557	1.49993	1.50566	1.51028	1.55355	1.55160	1.59258	1.58350	1.59746	1.59523	1.68585
1.7	1.49450	1.49862	1.50455	1.50896	1.55244	1.55040	1.59132	1.58239	1.59629	1.59405	1.68452
1.8	1.49342	1.49732	1.50340	1.50758	1.55133	1.54917	1.59007	1.58128	1.59515	1.59284	1.68315
1.9	1.49221	1.49592	1.50222	1.50613	1.55016	1.54785	1.58870	1.58008	1.59396	1.59155	1.68181
2.0	1.49094	1.49446	1.50096	1.50460	1.54896	1.54647	1.58732	1.57887	1.59268	1.59026	1.68040
2.1	1.48968	1.49291	1.49965	1.50303	1.54767	1.54506	1.58588	1.57761	1.59140	1.58893	1.67902
2.2	1.48836	1.49132	1.49831	1.50141	1.54634	1.54364	1.58438	1.57636	1.59001	1.58753	1.67763
2.3	1.48697	1.48959	1.49693	1.49967	1.54497	1.54216	1.58276	1.57502	1.58865	1.58612	1.67618
2.4	1.48549	1.48783	1.49546	1.49778	1.54348	1.54059	1.58107	1.57357	1.58714	1.58465	1.67474
2.5	1.48394	1.48597	1.49393	1.49588	1.54197	1.53901	1.57931	1.57218	1.58566	1.58314	1.67315
2.6	1.48234	1.48384	1.49226	1.49378	1.54037	1.53727	1.57747	1.57066	1.58403	1.58147	1.67149

No means for thermostatic control of the prism were available, and consequently settings could be made only while the room temperature remained constant within about 1°C.

When the temperature was found to have changed by more than a degree, the rocksalt prism was rotated back to its original position for the neon yellow line, the neon lamp was placed before the slit, and the fourth concave mirror was turned slightly by its levelling screws to bring the line back into its correct position in the middle of the exit slit. Since the temperature coefficient of *dispersion* was very small, this procedure sufficed for temperature changes of many degrees to either side of the original temperature for which the angles of incidence had been computed.

The dispersion data of rocksalt felt to be most reliable were those of Langley,[13] which were shown by him to fit the following Ketteler-Drude dispersion formula (at 20°C):

$$n^2=5.174714+0.0183744/(\lambda^2-0.015841)$$
$$-8949.52/(3145.695-\lambda^2).$$

The prism angle was measured on an ordinary spectrometer, and then by means of this formula, the angle of incidence was computed for each of

[13] Reference 12, p. 235.

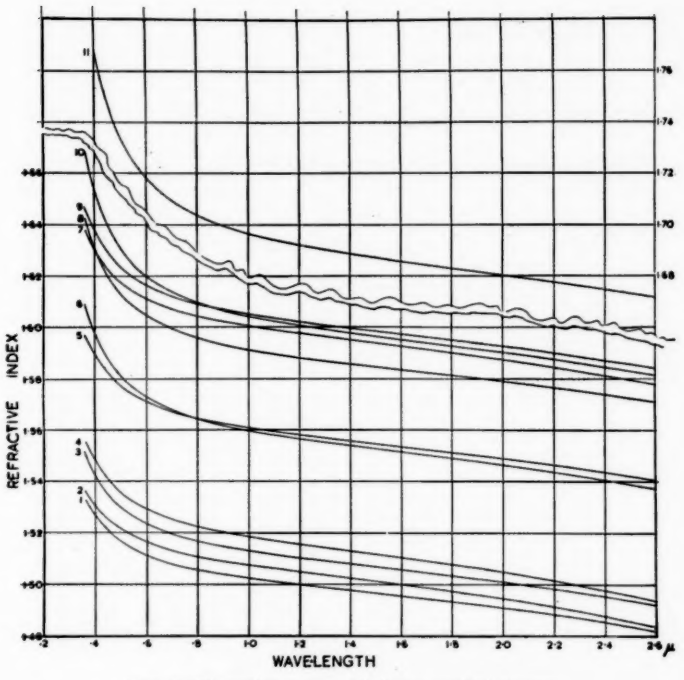


FIG. 2. Dispersion curves of eleven optical glasses.

the wave-lengths 0.60, 0.65, 0.70—up to 3μ. The instrumental calibration points so obtained fell on a very smooth curve which ran exactly through the 1.0140μ and the 0.7699μ points of mercury and potassium, respectively, previously obtained by direct observation.

Unfortunately the rocksalt calibration failed rather badly beyond 2μ on account of the very low dispersion of rocksalt in that region. The experimental procedure in making this calibration was to set the rocksalt prism at its proper angle corresponding to a desired wave-length, and then turn the first drum until the galvanometer showed a peak of deflection. However, on account of the low dispersion of rocksalt, a large range of wave-lengths could be covered by the glass prism before any appreciable change in the galvanometer deflection could be detected. Hence the rocksalt calibration had to be confined to the region below 2μ.

The construction of the instrument was found to be insufficiently rigid to hold its calibration for long periods, making it necessary to return repeatedly to the neon yellow line during observations. However, with care, good repetition of all readings was found to be quite possible. When using the thermopile, frequent returns to the well established 1.0140 mercury point were also made as a check on instrumental variations.

(b) *Quartz prism:* To extend the calibration curve beyond 2μ, a quartz Cornu prism was used instead of the rocksalt prism.

The dispersion of quartz has been found at the Physikalische Technische Reichsanstalt[14] to be accurately represented by the formula:

$$n^2=3.5968913+0.01064379/(\lambda^2-0.0106291)-138.20519/(111.45202-\lambda^2).$$

The curve obtained from this equation passes well through the data of Paschen,[10] but all Carvallo's points[15] fall slightly below it.

The instrument was therefore calibrated with the quartz prism, firstly for the 1.0140μ line as a check, which showed perfect agreement with the values obtained both by the direct mercury line and also indirectly from the rocksalt prism. The

[14] Zeits. f. Instrumentenk. **48**, 275 (1928).
[15] Carvallo, Comptes rendus **126**, 728 (1898).

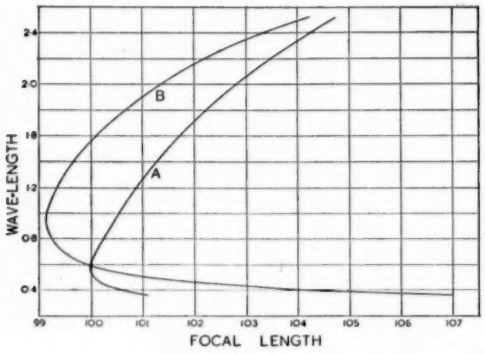


FIG. 3. Color curves for visual and photographic achromatism.

calibration was then taken from 1.8μ to the limit of transmission at 2.7μ, the quartz curve forming a perfect continuation of the rocksalt curve.

ACCURACY OF CALIBRATION CURVE

This is hard to estimate as it depends both on the accuracy of the data, and on instrumental errors. So far as self-consistency is a guide, all the observed points fit the smooth curve within about 0.004μ on the wave-length scale. Hence, taking into account the additional precision obtained by smoothing, we may say that the wave-length is probably accurate to within 0.003μ.

INFRARED INDICES OF OPTICAL GLASSES

With the kind assistance of the Bausch and Lomb Optical Company, 60° prisms of eleven different types of optical glass were obtained, and measured with the instrument. The results are given in Table II. At the top of each column[16] are given the usual visual indices, obtained with the aid of a laboratory spectrometer reading to 5 seconds, the indices being probably accurate to 2 or 3 in the fifth decimal place. The other indices were obtained with the infrared refractometer, and may be correct to 5 or 6 in the fifth decimal as this is the order of their self-consistency. So far as errors in the wave-length scale are concerned, an error of 0.003 in wave-length leads to an error of only about 1 in the fifth place of index.

These indices are shown plotted together in Fig. 2. It is noticeable that all the curves have an inflection at around 1.4 to 1.6μ, and that the dispersions of all the glasses in this region are approximately equal. The curves resemble closely those obtained by Rubens[5] and Abbott.[17]

INFRARED COLOR CURVES OF A THIN ACHROMATIC OBJECTIVE

The color curve of a thin achromatic doublet lens was calculated from the usual formula $1/f=\sum(n-1)c$ where f is the focal length and c the "net curvature" $(1/r_1-1/r_2)$. In Fig. 3 are shown two such curves; in (a) the ratio of c_a/c_b, the net curvatures of the two lens elements, is equal to 2.02, and in (b) it is 1.60. The first value is that required to unite the C and F lines at a common focus, giving a minimum focus at 0.555μ. In the second case, the minimum has fallen at about 0.95μ. In both curves, the focal length for the D line is assumed to be 100. It is seen what drastic changes are necessary in a lens if it is required to be achromatized in the near infrared instead of in the visible region.

[16] Under "maker," B-L indicates Bausch and Lomb, and PM Parra-Mantois. The "type numbers" for the three PM glasses are melt numbers.

[17] Abbott and Fowle, Astrophys. J. **11**, 135 (1900).

Light Signals on Moving Bodies as Measured by Transported Rods and Clocks

HERBERT E. IVES
Bell Telephone Laboratories, New York, N. Y.

(Received April 30, 1937)

1. INTRODUCTION

IN an earlier paper[1] I have pointed out that the null result of the Michelson-Morley and Kennedy-Thorndyke experiments may be explained by the assumptions:

(1) Moving measuring rods are altered in length when in uniform motion in the ratio

$$[(1-v^2/c^2)^{\frac{1}{2}}]^{n+1} : 1$$

in the direction of motion;

$$[(1-v^2/c^2)^{\frac{1}{2}}]^{n} : 1$$

at right angles to the direction of motion, and (2) Clocks are altered in frequency when in uniform motion in the ratio

$$[(1-v^2/c^2)^{\frac{1}{2}}]^{1-n} : 1,$$

where v is the velocity of the rods or clocks through the luminiferous ether, and c is the velocity of light.

No experiments, such as the Michelson-Morley experiment, using to and fro transits of light signals, are competent to establish the value of n in the above expression.[2] In the present paper I discuss the problem of light signals on moving bodies, taking these alterations of length and frequency as the basic physical phenomena, and assuming the value $n=0$ in the above expressions. This corresponds to a contraction of length in the direction of motion, of value $(1-v^2/c^2)^{\frac{1}{2}} : 1$; no alteration in length at right angles to the direction of motion; and a reduction of clock frequency in the ratio $(1-v^2/c^2)^{\frac{1}{2}} : 1$. This choice of value of n is that made by Fitzgerald, Lorentz, and Larmor, based on speculations and experiments apart from the Michelson-Morley experiment.[3] Neither the contraction of dimensions, nor the alteration of clock rate, has been subjected to direct experimental test.

2. THE FRAME OF REFERENCE AND TOOLS OF MEASUREMENT

In this discussion we shall assume light signals to be moving in a fixed ether, which we identify as the seat of the pattern of radiant energy received from the fixed stars. This pattern is substantially constant in character, suffering as a whole periodic variations with respect to the earth, shown by the phenomenon of stellar aberration. The argument for the necessity and reality of this frame of reference will be amplified at the close of the paper.

Since we are to consider the effects of variations in the properties of rods and clocks due to their motion through the ether, we shall compare these with rods and clocks which we shall postulate as *unaffected by transport*. Whether such rods and clocks can be realized in practice is broadly not important for carrying out a general argument. Using such rods and clocks, we define the *velocity of light* as the ratio of the distance between two points fixed in the ether, between which a light signal travels, to the time taken by transit. The velocity of light so defined we assume to be a constant, which we designate by c.

We note at once that the velocity of light with reference to bodies not fixed in the ether will differ from c, and that the value as measured on moving bodies with rods and clocks affected by motion may in general be expected also to be different from c. It is our problem to determine how different, in the case of the specific behavior of rods and clocks we have assumed.

[1] "Graphical Exposition of the Michelson-Morley Experiment," J. Op. Soc. Am. **27**, 177 (1937).

[2] This was recognized by Lorentz, who describes the corresponding factor in his expressions as "the origin of all our difficulties" (*Theory of Electrons*, p. 219), and chooses the value corresponding to our $n=0$ by consideration of the Kaufmann and Bucherer experiments. Birkhoff (*Relativity and Modern Physics*, p. 34) concludes that "the decision . . . must be made solely on the basis of facts established by means of physical experiments" and (p. 55) "It appears that no experimental data are available." He then discusses an experiment (the "transverse Doppler effect") which would be decisive.

[3] Larmor, *Aether and Matter*, 1900. The characterization of the frequency change as a prediction of the theory of relativity (1905) is chronologically unjustifiable.

3. The Measured Distances and Differences of Clock Readings Between Two Points on a Uniformly Moving Body

The most general method of measuring a length, e.g., the distance between points a and b on a body, is to place a "standard" measuring rod alongside the body, and read the positions of the two points a and b on divisions marked on this rod. In carrying out this comparison we ordinarily make the two readings at two different times, which would be without objection if the system were stationary. If the measuring rod is moving past the body, which is the general case, our observations must be made *simultaneously*. If our clocks at the two points of observation were unaffected by their separation this would present no complications. In our present problem, however, two clocks set alike at a differ when moved apart, and it becomes at once necessary to find what is the absolute time interval between measurements made "simultaneously" as indicated by our clocks.

Let the observing platform be moving with the velocity v. The velocity of light c, we shall assume to be obtainable to any desired degree of approximation, by a procedure which will be made clear in the sequel. Now assume our clock moved from a to b with the absolute velocity $v+W$. The true time [time in our framework with our unvarying clocks] T, which it takes to move between a and b is given by the relation

$$(v+W)T=L_a(1-v^2/c^2)^{\frac{1}{2}}+vT,$$

from which

$$T=L_a(1-v^2/c^2)^{\frac{1}{2}}/W \qquad (1)$$

where L_a is the distance ab when the body is stationary, as measured by an unvarying measuring rod. Now the time as indicated by the clock on reaching b will be:

$$T_b=T\left[1-\frac{(v+W)^2}{c^2}\right]^{\frac{1}{2}}$$

$$=\frac{L_a(1-v^2/c^2)^{\frac{1}{2}}}{W}\left[1-\frac{(v+W)^2}{c^2}\right]^{\frac{1}{2}}. \qquad (2)$$

Meanwhile the clock at a will indicate, at the same instant,

$$T_a=T\left(1-\frac{v^2}{c^2}\right)^{\frac{1}{2}}=\frac{L_a(1-v^2/c^2)^{\frac{1}{2}}}{W}\left(1-\frac{v^2}{c^2}\right)^{\frac{1}{2}} \qquad (3)$$

and there will be, after the clock, moved to b, has come to rest and resumed its rate of $\nu_0(1-v^2/c^2)^{\frac{1}{2}}$ (or a clock at b has been set at the indication of the moved clock as it passes) *a difference of setting* of the clock at b with respect to that at a

$$\Delta S=\frac{L_a(1-v^2/c^2)^{\frac{1}{2}}}{W}\left[\left(1-\frac{v^2}{c^2}\right)^{\frac{1}{2}}-\left(1-\frac{(v+W)^2}{c^2}\right)^{\frac{1}{2}}\right], \qquad (4)$$

which is the interval between apparently simultaneous observations.

An observer at a, making an observation at a predetermined indicated time actually makes the observation earlier on his clock by this difference of settings, and this amount earlier, to be reduced to true time, is to be increased in the ratio $1/(1-v^2/c^2)^{\frac{1}{2}}$, since his clock runs slow in the ratio $(1-v^2/c^2)^{\frac{1}{2}}:1$. The true time elapsed between the apparently simultaneous observations at a and b is therefore

$$\Delta T=L_a/W[(1-v^2/c^2)^{\frac{1}{2}}-(1-(v+W)^2/c^2)^{\frac{1}{2}}]. \qquad (5)$$

Let us now take a standard measuring rod, longer than the distance ab, which has been marked off in equal divisions, while stationary. Let the rod be moved past ab at an absolute velocity $v+Y$. The length of rod intercepted by ab is then

$$L_i=L_a(1-v^2/c^2)^{\frac{1}{2}}-Y\Delta T$$

or

$$L_i=L_a[(1-v^2/c^2)^{\frac{1}{2}}-(Y/W)\{(1-v^2/c^2)^{\frac{1}{2}}-(1-(v+W)^2/c^2)^{\frac{1}{2}}\}], \qquad (6)$$

which since the moving rod is contracted in the ratio $(1-(v+Y)^2/c^2)^{\frac{1}{2}}:1$, will be read as

$$L_i'=\frac{L_a\left[\left(1-\frac{v^2}{c^2}\right)^{\frac{1}{2}}-\frac{Y}{W}\left\{\left(1-\frac{v^2}{c^2}\right)^{\frac{1}{2}}-\left(1-\frac{(v+W)^2}{c^2}\right)^{\frac{1}{2}}\right\}\right]}{\left(1-\frac{(v+Y)^2}{c^2}\right)^{\frac{1}{2}}}. \qquad (7)$$

We next express the absolute velocities Y and W of measuring rod and clock in terms of *observable*

quantities, namely, their velocities as evaluated by divisions on the rod passing an observer on a clock per unit of time by the clock. In the case of the rod we imagine an observer at a, counting divisions against the "ticks" of his clock, whose rate is slow on the proportion $(1-v^2/c^2)^{\frac{1}{2}} : 1$. The rod being contracted in the ratio $(1-(v+Y)^2/c^2)^{\frac{1}{2}} : 1$, its *observed* velocity Y_0 is

$$Y_0=\frac{Y}{(1-v^2/c^2)^{\frac{1}{2}}(1-(v+Y)^2/c^2)^{\frac{1}{2}}}, \tag{8}$$

$$\text{from which } Y=\frac{Y_0(1-v^2/c^2)}{Y_0v/c^2+(1+Y_0^2/c^2)^{\frac{1}{2}}}. \tag{9}$$

For the moving clock, whose velocity relative to the body is W, we imagine an observer, on the clock, noting the time necessary for the light path ab, which he takes as his unit of length, to be traversed.[4] He thus observes his velocity increased over his true velocity, inversely as the length assumed by ab in motion, and inversely as his own clock rate, or

$$W_0=\frac{W}{(1-v^2/c^2)^{\frac{1}{2}}(1-(v+W)^2/c^2)^{\frac{1}{2}}}, \tag{10}$$

$$\text{from which } W=\frac{W_0(1-v^2/c^2)}{W_0v/c^2+(1+W_0^2/c^2)^{\frac{1}{2}}}. \tag{11}$$

We now have all the factors necessary to find the *observed* traversed distances and times of transit of light signals passing between the points a and b on a body moving with absolute velocity v.

4. Light Signals on a Single Moving Body

Consider first the apparent distance traversed by a light signal going from a to b, on the body moving with the uniform velocity v. This is the distance L_i' above. Substituting in it our values for W and Y in terms of the velocities as observed, W_0 and Y_0, we get

$$L_i'=L_s\frac{(1-v^2/c^2)^{\frac{1}{2}}-\dfrac{\dfrac{Y_0(1-v^2/c^2)}{Y_0v/c^2+(1+Y_0^2/c^2)^{\frac{1}{2}}}}{\dfrac{W_0(1-v^2/c^2)}{W_0v/c^2+(1+W_0^2/c^2)^{\frac{1}{2}}}}\left[(1-v^2/c^2)^{\frac{1}{2}}-\left\{1-\frac{\left(v+\dfrac{W_0(1-v^2/c^2)}{W_0v/c^2+(1+W_0^2/c^2)^{\frac{1}{2}}}\right)^2}{c^2}\right\}^{\frac{1}{2}}\right]}{\left[1+\left(\dfrac{v+\dfrac{Y_0(1-v^2/c^2)}{Y_0v/c^2+(1+Y_0^2/c^2)^{\frac{1}{2}}}}{c^2}\right)^2\right]^{\frac{1}{2}}}. \tag{12}$$

The solution of this equation[4] is

$$L_i'=L_s[(1+Y_0^2/c^2)^{\frac{1}{2}}+(Y_0/W_0)(1-(1+W_0^2/c^2)^{\frac{1}{2}})]. \tag{13}$$

On inspecting this result we note at once the significant fact that it is *independent of* v; the read value of the distance traversed by the light signal is a function of the stationary value L_s, of the distance ab, of the velocity of light c, and the measured velocities of rods and clocks, obtained in the manner described. Further, on expanding $(1+W_0^2/c^2)^{\frac{1}{2}}$ we get

$$L_i'=L_s[(1+Y_0^2/c^2)^{\frac{1}{2}}+Y_0W_0/2c^2], \tag{14}$$

which for $Y_0=0$ gives

$$L_i'=L_s. \tag{15}$$

Consider next the observed time of transit of the light signal. The true time is

$$T_t=\frac{L_s(1-v^2/c^2)^{\frac{1}{2}}}{c-v}, \tag{16}$$

which will be read by any stationary clock on the

[4] If the observer is called upon to insure that the clock is moving with uniform velocity with respect to the body (which is moving uniformly by assumption) he must watch the passage of uniformly spaced divisions past the clock. These divisions can be those on our standard measuring rod whose uniform velocity with respect to the body (which might be zero) can be assured by independent observation, with the clock held stationary on the body. This procedure avoids any commitment as to the methods of length measurement on the moving body.

body as

$$T_t=\frac{L_s(1-v^2/c^2)^{\frac{1}{2}}(1-v^2/c^2)^{\frac{1}{2}}}{c-v}=(L_s/c)(1+v/c). \tag{17}$$

Now the clock at b is set back, due to its journey from a by the quantity given above (4). Substituting in this W_0, the observed value of W, we get

$$T_t'=\frac{L_s}{c}\left(1+\frac{v}{c}\right)-\frac{L_s(1-v^2/c^2)^{\frac{1}{2}}}{\dfrac{W_0(1-v^2/c^2)}{W_0v/c^2+(1+W_0^2/c^2)^{\frac{1}{2}}}}\left[\left(1-\frac{v^2}{c^2}\right)^{\frac{1}{2}}-\left(1-\frac{\left(v+\dfrac{W_0(1-v^2/c^2)}{W_0v/c^2+(1+W_0^2/c^2)^{\frac{1}{2}}}\right)^2}{c^2}\right)^{\frac{1}{2}}\right]. \tag{18}$$

The solution of this[5] is

$$T_t'=\frac{L_s}{c}\left[\frac{1+W_0/c-(1+W_0^2/c^2)^{\frac{1}{2}}}{W_0/c}\right]. \tag{19}$$

Inspecting this we note once more that our value is *independent of* v as long as we keep the observed velocities of rods and clocks constant. (The absolute values to meet these conditions of course vary with every value of v.) Further, on expanding $(1+W_0^2/c^2)^{\frac{1}{2}}$, we find, if W_0 is small

$$T_t'=(L_s/c)(1-W_0/2c), \tag{20}$$

so that for infinitesimal velocities of the clock, the read time approaches L_s/c.

Next consider the *velocity of light* as measured by sending a light signal between two points on a moving body such as the earth. The observed or measured value will be

$$\frac{L_t'}{T_t'}=\frac{L_s}{L_s/c}\,\frac{[(1+Y_0^2/c^2)^{\frac{1}{2}}+Y_0/W_0(1-(1+W_0^2/c^2)^{\frac{1}{2}})]}{\left[\dfrac{1+W_0/c-(1+W_0^2/c^2)^{\frac{1}{2}}}{W_0/c}\right]} \tag{21}$$

$$=c\,\mathfrak{L}/\mathfrak{T}=c_m, \tag{22}$$

using $\mathfrak{L}$ and $\mathfrak{T}$ for the measured quantities in brackets and c_m for the velocity of light measured with transported rods and clocks, *which is different from* c, our fundamental constant. Since $\mathfrak{L}$ and $\mathfrak{T}$ are functions of c, it will be possible to express c entirely in time of measured velocities of clocks, rods and light. From the preceding work it follows that for $Y_0=0$, *the value c is approximated as the velocity of the rods and clocks is indefinitely reduced.*

A further observation on the absolute velocity of light, which we promised above would be shown to be a determinable quantity. From the foregoing discussion it follows that the value c can be obtained to any desired degree of approximation, by sending a light signal to a point on a moving body, whose distance is obtained by the use of measuring rods, and determining its time of arrival by a clock carried to that point, provided the rods and clocks used are moved at negligible velocities with respect to the moving body. The method used in all practical determinations of the velocity of light meets these requirements. The distance is measured by laying rods end to end, each rod being stationary on the body when the measurement is recorded; the time is measured by a clock stationary on the body, by virtue of the invariable procedure of sending the light signal to a mirror and back to the origin.

5. Light Signals on Several Moving Bodies

We now investigate the measured values of time and distance for a light signal traversing several moving bodies simultaneously. In Fig. 1, let B_0, B_1, B_2, etc. represent the several bodies in motion, B_0 being the luminiferous ether, B_1, B_2,

[5] See Appendix. I am indebted to Mr. T. C. Fry for the formal solutions.

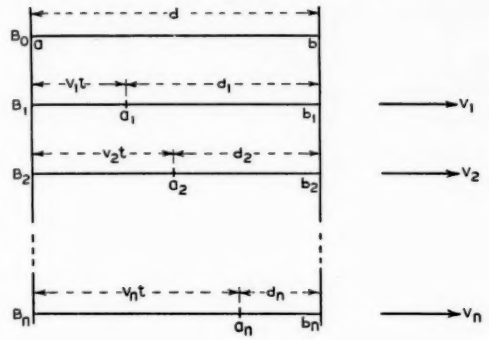


FIG. 1.

etc., being considered for purposes of visualization as moving bands all in immediate contact with each other, and with the light source positioned at a. Under these conditions, the light signal when emitted can be instanteously and simultaneously impressed, say upon a photosensitive surface, upon each moving band; and when received, as by a photoelectric cell at b, its reception may be instantaneously and simultaneously recorded on each moving band.

Let the distance traversed by the light signal in the ether (B_0) be designated as d, and let d_1, d_2, etc., be the distances between the initial and final traces of the light signals (b_1-a_1, b_2-a_2) on the several moving bands. Let the time of transit of the light signal from a to b be

$$t=d/c. \tag{23}$$

Then the exact physical statement of the event on each band is

$$\begin{aligned} d_1&=d(1-v_1/c),\\ d_2&=d(1-v_2/c),\\ &\cdots\cdots\\ d_n&=d(1-v_n/c). \end{aligned} \tag{24}$$

Now the trace $d(1-v_1/c)$ on the band moving with the velocity v_1 would, if the band were stopped, taking the value of the contraction as $(1-v_1^2/c^2)^{\frac{1}{2}}$ in the direction of motion, assume the value

$$\frac{d(1-v_1/d)^{\frac{1}{2}}}{(1-v_1^2/c^2)^{\frac{1}{2}}}. \tag{25}$$

Now this is merely the L_s of our previous section. Therefore, to get the measured values, using our moved rods and clocks we have

$$\begin{aligned} d_1'&=\frac{d(1-v_1/c)}{(1-v_1^2/c^2)^{\frac{1}{2}}}[(1+Y_0^2/c^2)^{\frac{1}{2}}+(Y_0/W_0)\{1-(1+W_0^2/c^2)^{\frac{1}{2}}\}],\\ d_2'&=\frac{d(1-v_2/c)}{(1-v_2^2/c^2)^{\frac{1}{2}}}[(1+Y_0^2/c^2)^{\frac{1}{2}}+(Y_0/W_0)\{1-(1+W_0^2/c^2)^{\frac{1}{2}}\}],\\ d_n'&=\frac{d(1-v_n/c)}{(1-v_n^2/c^2)^{\frac{1}{2}}}[(1+Y_0^2/c^2)^{\frac{1}{2}}+(Y_0/W_0)\{1-(1+W_0^2/c^2)^{\frac{1}{2}}\}]. \end{aligned} \tag{26}$$

These are different for each moving band.

Also from the previous section we have that the observed time of transit is $L_s/c\cdot\mathcal{T}$, so that we have

$$\begin{aligned} t_1'&=\frac{d}{c}\frac{(1-v_1/c)}{(1-v_1^2/c^2)^{\frac{1}{2}}}\left[\frac{1+W_0/c-(1+W_0^2/c^2)^{\frac{1}{2}}}{W_0/c}\right]\\ &=\frac{t(1-v_1/c)}{(1-v_1^2/c^2)^{\frac{1}{2}}}\left[\frac{1+W_0/c-(1+W_0^2/c^2)^{\frac{1}{2}}}{W_0/c}\right],\\ t_2'&=\frac{t(1-v_2/c)}{(1-v_2^2/c^2)^{\frac{1}{2}}}\left[\frac{1+W_0/c-(1+W_0^2/c^2)^{\frac{1}{2}}}{W_0/c}\right],\\ t_n'&=\frac{t(1-v_n/c)}{(1-v_n^2/c^2)^{\frac{1}{2}}}\left[\frac{1+W_0/c-(1+W_0^2/c^2)^{\frac{1}{2}}}{W_0/c}\right]. \end{aligned} \tag{27}$$

These values are different for each band.

The *observed velocity of light* on each band, provided we have adhered to the same observed manipulations of rods and clocks is

$$\frac{d_1'}{t_1'}=\frac{d_2'}{t_2'}=\frac{d_n'}{t_n'}=c\frac{[(1+Y_0^2/c^2)^{\frac{1}{2}}+(Y_0/W_0)\{1-(1+W_0^2/c^2)^{\frac{1}{2}}\}]}{\left[\dfrac{1+W_0/c-(1+W_0^2/c^2)^{\frac{1}{2}}}{W_0/c}\right]}, \tag{28}$$

which for zero rod velocity and small clock velocity, approximates c.

An observation which is obvious is that our Eqs. (26), (27) for the case of infinitesimal velocities of rods and clocks are (for the direction of motion, which has alone been presented) the well-known Lorentz transformations,[6] for

$$d_1'=\frac{d(1-v_1/c)}{(1-v_1^2/c^2)^{\frac{1}{2}}}=\frac{d-v_1(d/c)}{(1-v_1^2/c^2)^{\frac{1}{2}}}=\frac{d-v_1t}{(1-v_1^2/c^2)^{\frac{1}{2}}}, \tag{29}$$

which is expressed by Lorentz as

$$x'=\frac{x-vt}{(1-v^2/c^2)^{\frac{1}{2}}} \tag{30}$$

and

$$t_1'=\frac{t(1-v_1/c)}{(1-v_1^2/c^2)^{\frac{1}{2}}}=\frac{t-(v_1/c)t}{(1-v_1^2/c^2)^{\frac{1}{2}}}=\frac{t-v_1d/c^2}{(1-v_1^2/c^2)^{\frac{1}{2}}}, \tag{31}$$

which is expressed by Lorentz as

$$t'=\frac{t-vx/c^2}{(1-v^2/c^2)^{\frac{1}{2}}}, \tag{32}$$

which he calls "local time." Since all the clocks (Fig. 1) are recording an event (arrival of the light signal) at *the same place, at the same time,* each giving a *different* value, the term "local time" is somewhat of a misnomer for the array of indications possible. "Clock readings" is a more accurately descriptive term.

We now develop an interesting consequence of the contractions we have assumed as the basis of our discussion, namely, that all the phenomena of measurement of light paths and times of transit may be expressed entirely in terms of the *measured* velocities of the bodies and of light.

Let us take any pair of read distances, such as d_1' and d_2' and solve for one in terms of the other.

$$\frac{d_2'}{d_1'}=\frac{d(1-v_1/c)}{d(1-v_2/c)}\left[\frac{1-v_2^2/c^2}{1-v_1^2/c^2}\right]^{\frac{1}{2}}, \tag{33}$$

from which

$$d_2'=d_1'\left[\frac{1-\dfrac{(v_2-v_1)}{c(1-v_1v_2/c^2)}}{\left\{1-\dfrac{(v_2-v_1)^2}{c^2(1-v_1v_2/c^2)}\right\}}\right], \tag{34}$$

with similar expressions for the measured times.

We note that this equation is similar in form to the expressions from which it is derived, the quantity $(v_2-v_1)/(1-v_1v_2/c^2)$ taking the place of v_1 and v_2. Now v_2-v_1 is an absolute difference of velocities, which suggests that we investigate the corresponding term for observed or measured differences of velocity.

Consider a point on a uniformly moving body of absolute velocity v_2, which is passing our observation platform moving at the velocity v_1, with the difference of velocity v_2-v_1. Now the absolute distance between two marks a, b, on our body is $L_s(1-v_1^2/c^2)^{\frac{1}{2}}$ so that the true time for the transit of the moving point from a to b is

$$L_s(1-v^2/c^2)^{\frac{1}{2}}/(v_2-v_1), \tag{35}$$

[6] In the procedure used to obtain the Lorentz transformations by starting with the postulate that the velocity of light is always c, is the concealed restriction that all clocks and rods are to be manipulated at infinitesimal speeds. The final results are subject to this same restriction, which when recognized leads to such dicta as "clocks must be excluded from time measurement," and that the times recorded by a clock are "incorrect." In view of the role played in both macroscopic and atomic physics (e.g., spectroscopy) by rods and clocks, no system which throttles them can be physically satisfactory.

which will be read by a clock on the body moving at velocity v_1, as

$$L_s(1-v^2/c^2)/(v_2-v_1). \tag{36}$$

Now from this is to be subtracted the change in clock setting due to the transport of the clock from a to b, giving

$$T'=\frac{L_s(1-v^2/c^2)}{v_2-v_1}-\frac{L_s(1-v^2/c^2)^{\frac{1}{2}}}{W}\left[\left(1-\frac{v^2}{c^2}\right)^{\frac{1}{2}}-\left(1-\frac{(v+W)^2}{c^2}\right)^{\frac{1}{2}}\right]. \tag{37}$$

Substituting W_0 from (11) and solving we get

$$T'=\frac{L_s(1-v_1v_2/c^2)}{v_2-v_1}+\frac{L_s}{cW_0}\left[1-\left(1+\frac{W_0^2}{c^2}\right)^{\frac{1}{2}}\right], \tag{38}$$

from which, putting v_{2-1} for the observed velocity difference,

$$v_{2-1}=\frac{L_s}{T}=\frac{1}{\dfrac{1-v_1v_2/c^2}{v_2-v_1}+\dfrac{1}{cW_0}\left[1-\left(1+\dfrac{W_0^2}{c^2}\right)^{\frac{1}{2}}\right]}, \tag{39}$$

giving

$$\frac{v_2-v_1}{1-v_1v_2/c^2}=\frac{v_{2-1}}{1-v_{2-1}\left[1-(1+W_0^2/c^2)^{\frac{1}{2}}\right]}. \tag{40}$$

If now we substitute this value for $(v_2-v_1)/(1-v_1v_2/c^2)$ in the relation beween d_1' and d_2', and also substitute for c its value in terms of c_m we finally have the relation beween the measured values d_1' and d_2' *entirely in terms of the observed velocities of rods, clocks, and light.* Therefore, as long as we perform our measurements on the distances and times of transit of light signals, with rods and clocks which we move at the same measured velocities, we can give a complete description without reference to our absolute framework.

Returning to the expression

$$(v_2-v_1)/(1-v_1v_2/c^2),$$

note that for $W_0 \leqq 0$ this becomes simply

$$\frac{v_2-v_1}{1-v_1v_2/c^2}=v_{2-1}. \tag{41}$$

Inserting this in the expression for d_2' in terms of d_1' we get

$$d_2'=\frac{d_1'(1-v_{2-1}/c)}{(1-(v_{2-1})^2/c^2)^{\frac{1}{2}}}. \tag{42}$$

Since c is also the measured value of the velocity of light for $W_0 \leqq 0$, $Y_0=0$, we see that we have for this case an expression exactly similar in form to the original expressions (26)

$$d_1'=\frac{d(1-v/c)}{(1-v^2/c^2)^{\frac{1}{2}}}$$

in terms of absolute velocities and with rods and clocks moved infinitely slowly. It is probable that the expressions for d_2' in terms of d_1' for finite values of W_0 and Y_0 would take similar forms after the performance of the proper transformations.

6. Summary of Results

The postulates used in this discussion may be arranged and listed as follows:

A postulate covering the velocity of light and its measurement

I. The velocity of light in the luminiferous ether, when measured by light signals between two points at rest in the ether, by clocks and rods unaffected by transport, is a constant c.

A postulate covering the behavior of material rods and clocks

II. The lengths of material rods (in the direction of motion) and the frequencies of material clocks, are reduced by the factor $(1-v^2/c^2)^{\frac{1}{2}}$, where v is the velocity of the rod or clock, and c the velocity of light, measured as in I.

By the use of these postulates we *conclude* that

III. Measurements of light signals made on single moving bodies yield results which can be expressed entirely in terms of the measured velocities of rods, clocks, and light, and are identical whatever the velocities of the moving bodies, so long as the same measured velocities of rods and clocks are adhered to.

Using III in conjunction with I and II we find expressions for light signals traversing several moving bodies.[7] Principal conclusions are:

(1) Measured distances and times of transit of identical light signals are different for each moving body.

(2) The velocity of light measures the same on each body, the actual value being a function of the observed velocities of the measuring rods and clocks used, which approximates c as those velocities become small compared to the velocity of light.

(3) The expressions found approximate to the Lorentz transformations as the velocities of rods and clocks become small compared to the velocity of light.

7. Arguments for the Framework and Measuring Methods

In the treatment of light signals on moving bodies here presented we have assumed the existence of a luminiferous ether, and have used throughout the conception of true or absolute time, such as could be established if clocks could be moved from place to place through the ether without being affected by their transport. Both these are conceptions which it has become common to deny as necessary, in view of the possible elimination of all terms representing these quantities in the final equations which we have been able to derive, in which occur only quantities determined by measurements involving the moving bodies. I give now some considerations which I think, support these conceptions in spite of the common argument against their necessity.

In the case of the time factor the objection is commonly stated in terms of the idea of simultaneity at separated points. If this cannot be determined, it is argued that it has "no meaning." Now the distinction which must be made here is between *nonexistence* and *indeterminacy*. It is now well recognized in other departments of physics that phenomena which are altered by the process of measurement become as a consequence *indeterminate*. In the case of simultaneity we have seen that the transport of a clock to a distant point is such a function of the velocity that it prevents the velocity being known. This in turn prevents the correlation of the clock indications at the two points. The *existence* of true simultaneity is however easily established by the following observation:

Consider three clocks A, B and B', all having the same rates when together, and let B and B' be moved to a distant point on the common moving platform. Send a series of equidistant pulses from B to A, and let A be set to be synchronous with the pulses it receives. Then let the pulses sent from A be used to set B'. We will then have B and B' going at the same rate, but with a difference of phase. Let us now set up a number of other clocks at B, B' all of the same rate, but at phases between B and B'. Now by taking a sufficiency of these clocks, *one will be* beating simultaneously with A, irrespective of the fact that we cannot determine which it is, within the limits set by B and B', which give a measure of the indeterminacy.

Einstein offers as a substitute for this indeterminacy the proposal to *define* simultaneity as given by *half the* time taken for a to and fro light signal. This of course will agree with the simultaneity above discussed only if the system is stationary in the light transmitting medium. It corresponds to "local time" when clocks and rods are moved infinitely slowly. As already pointed out this involves us in the predicament that parallel light signals traversing several bodies are, when their arrival at a terminal point is recorded on each moving body, reported as arriving at a *different* time on each body, although constituting a *single event at a single place*,—simultaneity by the most rigid criterion.

[7] It is possible, by assuming I and III, to derive II, by imagining light signals, sent over a to and fro path, to take an invariant measured time, and by requiring that the measured distance and time at the far point shall be invariant whatever the velocities of the rods and clocks used in measurement. Both this line of attack (assumption of invariance), and that followed in the paper (assumption as to behavior of material rods and clocks) requires I, which is thus characterized as the fundamental assumption in the treatment of light signals.

The definition of simultaneity proposed by Einstein is violated immediately clocks moving at finite velocities on the bodies are used to establish "local" times. The time of arrival of a light signal at a distant point is then a function of the clock speed, given by (27) which is no longer half the interval recorded by a clock at the origin for a to and fro signal. These various differing measures of the time of arrival of a light signal at a distant point are simply manifestations of the dependence of our results on the characteristics of the tools of measurement used. They none of them give us the simultaneity above discussed, which could be found only with clocks and rods unaffected by motion, which have been here taken as the basis for discussing material clocks and rods which *are* affected.

We are saved from any *practical* difficulties in connection with distant synchronization by the extremely low speeds at which clocks and rods can be moved.

Turning next to the question of the ether, we admit at once its elusiveness by any measurements of the sort we have been considering (based on the value $n=0$ in the contraction expressions), for the result of the length and clock rate contractions is to make these measurements invariant with v, the velocity of the system with respect to the ether. Referring however to a preceding paper,[1] it was there stressed that an uncompensated difference in the character of the light signal for different velocities of the system remained, namely the *direction* of the pulse or wave front, which enters or leaves the interferometer at the *aberration angle*. This is a different characteristic of the light from the measurements of distances and times of transit to which we have directed attention here, and it is in this directional factor that we have the strongest evidence for the existence of the luminiferous ether.

This evidence we get from the *aberration of light from the stars*. The earth swims in a sea of radiant energy, resolvable into a pattern of wave fronts maintaining over long periods substantial constancy of relation in direction to each other. In the course of a year this pattern of radiant energy experiences, *as a whole*, a to and fro oscillation of direction. The interpretation of Bradley, that it is the earth which is moving in an orbit, through the substantially unchanging pattern of light waves, is as compelling today as ever. The alternative, that this entire energy pattern, together with the stellar universe with which it is associated (either immovably or with a steady state of drift, which alternative being indeterminate) oscillates back and forth yearly, is improbable in the ratio of the number of stars to the number of the observing platforms, namely one. The evidence is thus not metric, but by sheer weight of enormous probability. It is the same evidence that leads us to accept the Copernican as against the Ptolemaic system.[8]

It must not be forgotten in the discussion of this subject that the Michelson-Morley experiment, which is the original and crucial problem, only demands invariance of light signals with the velocity of the moving platform of measurement *on the premise that the earth is moving*,—there is no other motion involved in the experiment. If this is not agreed to then the null result proves nothing with regard to invariance, and the whole discussion is futile. Now the annual motion of the earth (which is the motion it is sought to detect in the Michelson-Morley experiment) we deduce from the energy pattern reaching us from outer space, in other words we observe that the earth moves with respect to this energy pattern. The Michelson-Morley experiment is performed on the assumption that a light signal initiated in the apparatus becomes a part of this energy pattern with respect to which the earth moves. The alternative would be that the signal partakes of the velocity of the source, which again would nullify the whole discussion. Calling upon the Michelson-Morley experiment as proof of the invariance of light signal phenomena therefore carries with it the acceptance of the luminiferous ether here assumed.

[8] The relativity treatments of stellar aberration all start with a "wave front" from a star, thus of necessity recognizing the pattern of radiant energy here considered as a manifestation of the ether. The common relativity treatment of aberration concerns itself with a broad section of a wave front, and shows that, due to the "local" time difference at points along the receiving surface the wave front will always be interpreted as being normal to the direction of propagation of the wave front section as a whole, with respect to the earth. This does not touch the question of the cause of the direction taken by the wave section or ray, and is incidentally not the case of a star image formed by a telescope. The second-order term in the "relativity" expression for aberration follows at once along the lines here followed, from the contraction of the measuring plane in the moving telescope.

8. Conclusion

In conclusion I wish to emphasize the three factors of this treatment which I believe to be most important. First, is the exact specification of the framework, and the characteristics of the tools of measurement and their use. Second, is the choice of the length and frequency contractions as the basic phenomena in the behavior of moving bodies. Third, is the admission of velocities of the measuring instruments other than infinitesimal, with the consequent recognition that the measured velocity of light is a function of the measuring conditions, and can have other values than the constant c.

Appendix

Solutions of equations involving W, W_0, Y and Y_0

(1) W and Y in terms of W_0 and Y_0.
Let $v/c=x$, $W/c=z$, $W_0/c=\alpha$.

(10) is then
$$\alpha=\frac{z}{(1-x^2)^{\frac{1}{2}}(1-(x+z)^2)^{\frac{1}{2}}},$$
$$z^2/\alpha^2=(1-x^2)(1-[x+z]^2)$$
$$=(1-x^2)(1-x^2-2xz-z^2),$$
$$z^2[1+\alpha^2(1-x^2)]+2zx\alpha^2(1-x^2)-\alpha^2(1-x^2)^2=0,$$
$$z=\frac{-x\alpha^2(1-x^2)\pm\{x^2\alpha^4(1-x^2)^2+\alpha^2(1-x^2)^2[1+\alpha^2(1-x^2)]\}^{\frac{1}{2}}}{1+\alpha^2(1-x^2)}$$
$$=\frac{\alpha(1-x^2)[-x\alpha\pm(1+\alpha^2)^{\frac{1}{2}}]}{((1+\alpha)^{\frac{1}{2}}-x\alpha)((1+\alpha)^{\frac{1}{2}}+x\alpha)},$$

taking the positive sign
$$=\frac{\alpha(1-x^2)}{x\alpha+(1+\alpha^2)^{\frac{1}{2}}}$$

or
$$\frac{W}{c}=\frac{W_0/c(1-v^2/c^2)}{W_0v/c^2+(1+W_0^2/c^2)^{\frac{1}{2}}}.$$

The same method of solution gives a similar expression for Y in terms of Y_0.

(2) Solution of
$$L_i'=\frac{L_s\{(1-v^2/c^2)^{\frac{1}{2}}-(Y/W)[(1-v^2/c^2)^{\frac{1}{2}}-(1-(v+W)^2/c^2)^{\frac{1}{2}}]\}}{(1-(v+Y)^2/c^2)^{\frac{1}{2}}},$$

adding the symbols y for Y/c and β for Y_0/c we have
$$L_i'=\frac{L_s[(1-x^2)^{\frac{1}{2}}-(y/z)\{(1-x^2)^{\frac{1}{2}}-(1-(x+z)^2)^{\frac{1}{2}}\}]}{(1-(x+y)^2)^{\frac{1}{2}}}.$$

Now
$$x+z=x+\frac{\alpha(1-x^2)}{x\alpha+(1+\alpha^2)^{\frac{1}{2}}}=\frac{\alpha+x(1+\alpha^2)^{\frac{1}{2}}}{x\alpha+(1+\alpha^2)^{\frac{1}{2}}},$$

from which
$$(1-(x+z)^2)^{\frac{1}{2}}=\frac{(1-x^2)^{\frac{1}{2}}}{x\alpha+(1+\alpha^2)^{\frac{1}{2}}}$$

and
$$(1-(x+y)^2)^{\frac{1}{2}}=\frac{(1-x^2)^{\frac{1}{2}}}{x\beta+(1+\beta^2)^{\frac{1}{2}}}$$

so that
$$L_i'=\frac{L_s[(1-x^2)^{\frac{1}{2}}-y/z\{(1-x^2)^{\frac{1}{2}}-(1-x^2)^{\frac{1}{2}}/(x\alpha+(1+\alpha^2)^{\frac{1}{2}})\}]}{(1-x^2)^{\frac{1}{2}}/(x\beta+)1+\beta^2)^{\frac{1}{2}})}$$
$$=L_s\left\{1-\frac{\frac{\beta(1-x^2)}{x\beta+(1+\beta^2)^{\frac{1}{2}}}}{\frac{\alpha(1-x^2)}{x\alpha+(1+\alpha^2)^{\frac{1}{2}}}}\left[1-\frac{1}{x\alpha+(1+\alpha^2)^{\frac{1}{2}}}\right]\right\}(x\beta+(1+\beta^2)^{\frac{1}{2}})$$
$$=L_s[(1+\beta^2)^{\frac{1}{2}}+(\beta/\alpha)(1-(1+\alpha^2)^{\frac{1}{2}})]$$

or
$$L_i'=L_s[(1+Y_c^2/c^2)^{\frac{1}{2}}+(Y_0/W_0)\{1-(1+W_0^2/c^2)^{\frac{1}{2}}\}].$$

(3) Solution of
$$T_i'=L_s/c(1+v/c)-L/W(1-v^2/c^2)^{\frac{1}{2}}[(1-v^2/c^2)^{\frac{1}{2}}-(1-(x+W)^2/c^2)^{\frac{1}{2}}].$$

Making the substitutions indicated in (1) and (2)

$$T_i' = \frac{L_s}{c}(1+x) - \frac{L_s/c}{z}(1-x^2)^{\frac{1}{2}}[(1-x^2)^{\frac{1}{2}} - (1-(x+z)^2)^{\frac{1}{2}}]$$

$$= \frac{L_s}{c}(1+x) - \frac{L_s}{c}\,\frac{(1-x^2) - (1-x^2)^{\frac{1}{2}}(1-(x+z)^2)^{\frac{1}{2}}}{z},$$

putting $(1-(x+z)^2)^{\frac{1}{2}} = \dfrac{(1-x^2)^{\frac{1}{2}}}{x\alpha + (1+\alpha^2)^{\frac{1}{2}}}$ as above found, and substituting for z its value in terms of α

$$T_i' = \frac{L_s}{c}(1+x) - L_s(1-x^2)\left[\frac{1 - \dfrac{1}{x\alpha + (1+\alpha^2)^{\frac{1}{2}}}}{\dfrac{\alpha(1-x^2)}{x\alpha + (1+\alpha^2)^{\frac{1}{2}}}}\right]$$

$$= \frac{L_s}{c}\left[\frac{\alpha + 1 - (1+\alpha^2)^{\frac{1}{2}}}{\alpha}\right]$$

$$= \frac{L_s}{c}\left[\frac{W_0/c + 1 - (1 + W_0^2/c^2)^{\frac{1}{2}}}{W_0/c}\right].$$